NOV 2 - 1931

OCTOBER 1931

Journal of Educational Research

Registered June 4, 1929

CONTENTS

		PAGE
Needed Research in Diagnosing Personality and Conduct	*Percival M. Symonds*	175
The Arithmetical Backgrounds of Young Children	*Clifford Woody*	188
A Method for Measuring Retention in Reading	*Alvin C. Eurich*	202
Improving College Freshmen in Spelling	*Walter S. Guiler*	209
The Objectivity, Reliability, and Validity of an Essay Examination Graded by Rating	*Verner Martin Sims*	216
Shorter Articles		224
Editorials		231
Reviews		233
Research Abstracts and Bibliographies		239
Research News and Communications		246

THE JOURNAL OF EDUCATIONAL RESEARCH is indexed in the Education Index

---

Copyright 1931, by Public School Publishing Co.

Entered as second-class matter at the Post Office at Bloomington, Illinois, under the act of March 3, 1879

JOURNAL OF EDUCATIONAL RESEARCH

EDITORIAL BOARD

E. J. ASHBAUGH, *Dean, School of Education and Professor of Education, Miami University, Oxford, Ohio*

A. S. BARR, *Chairman, Editorial Board, Professor of Education, University of Wisconsin, Madison, Wisconsin*

CARTER V. GOOD *Professor of Education, University of Cincinnati, Cincinnati, Ohio*

ASSOCIATE EDITORS

FRED C. AYER, *Professor of Educational Administration, University of Texas, Austin, Texas*

HARRY J. BAKER, *Director, Psychological Clinic, Detroit Public Schools, Detroit, Michigan*

FOWLER D. BROOKS, *Head of Departments of Education and Psychology, DePauw University, Greencastle, Indiana*

B. R. BUCKINGHAM, *Lecturer in Education, Harvard University, Cambridge, Massachusetts*

W. W. CHARTERS, *Director, Bureau of Educational Research and Professor of Education, Ohio State University, Columbus, Ohio*

VIRGIL E. DICKSON, *Assistant Superintendent of Schools and Director of Bureau of Research and Guidance, Berkeley Public Schools, Berkeley, California*

CHARLES FORDYCE, *Assistant Director, Bureau of Educational Reference and Research, University of Nebraska, Lincoln, Nebraska.*

WILLIAM S. GRAY, *Dean, School of Education and Professor of Education, University of Chicago, Chicago, Illinois*

CHARLES E. GREENE, *Director of Research, Denver Public Schools, Denver, Colorado*

M. E. HAGGERTY, *Dean, College of Education and Professor of Education, University of Minnesota, Minneapolis, Minnesota*

V. A. C. HENMON, *Chairman, Department of Psychology and Professor of Psychology, University of Wisconsin, Madison, Wisconsin.*

KARL J. HOLZINGER, *Professor of Education, University of Chicago, Chicago, Illinois.*

ERNEST HORN, *Director, University Elementary School and Professor of Education, State University of Iowa, Iowa City, Iowa*

EARL HUDELSON, *Dean, College of Education and Professor of Education, West Virginia University, Morgantown, West Virginia*

ARTHUR W. KALLOM, *First Assistant Director, Department of Educational Investigation and Measurement, Boston Public Schools, Boston, Massachusetts*

FREDERICK J. KELLY, *Chief, Section of Colleges and Professional Schools, Office of Education, Department of the Interior, Washington, D. C.*

WILLIAM A. McCALL, *Professor of Education, Teachers College, Columbia University, New York City*

EUGENE A. NIFENECKER, *Director of Reference, Research, and Statistics, Board of Education, 500 Park Avenue, New York City*

W. J. OSBURN, *Director of Research, State Department of Education, and Professor of School Administration, Ohio State University, Columbus, Ohio*

SHELTON PHELPS, *Director of Instruction and Professor of Education, George Peabody College, Nashville, Tennessee*

DOUGLAS E. SCATES, *Director, Research and Statistics, Cincinnati Public Schools, Cincinnati, Ohio*

JESSE B. SEARS, *Professor of Education, Stanford University, California*

JOHN L. STENQUIST, *Director, Bureau of Educational Research, Department of Education and Lecturer, Johns Hopkins University, Baltimore, Maryland*

GEORGE D. STRAYER, *Director, Division of Field Studies, Institute of Educational Research, and Professor of Education, Teachers College, Columbia University, New York City*

PERCIVAL M. SYMONDS, *Associate Professor of Education, Teachers College, Columbia University, New York City*

W. W. THEISEN, *Assistant Superintendent of Schools, Milwaukee, Wisconsin*

M. R. TRABUE, *Executive Secretary of the Committee on Individual Diagnosis and Training, Institute for Employment Stabilization Research, University of Minnesota, Minneapolis, Minnesota*

W. L. UHL, *Dean, School of Education, and Professor of Education, University of Washington, Seattle, Washington.*

CARLETON W. WASHBURNE, *Superintendent of Schools, Winnetka Public Schools, Winnetka, Illinois*

GUY M. WHIPPLE, *Secretary, National Society for the Study of Education, Danvers, Massachusetts.*

CLIFFORD WOODY, *Director, Bureau of Educational Reference and Research and Professor of Education, University of Michigan, Ann Arbor, Michigan*

THE JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Company. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

Subscriptions and all communications about reprints or other business matters should be addressed to the PUBLIC SCHOOL PUBLISHING COMPANY, at Bloomington, Illinois.

Address all communications on editorial matters to A. S. Barr, Department of Education, University of Wisconsin, Madison, Wisconsin.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume XXIV OCTOBER 1931 Number 3

NEEDED RESEARCH IN DIAGNOSING PERSONALITY AND CONDUCT

PERCIVAL M. SYMONDS
Teachers College, Columbia University

THE measurement movement in recent years has broadened out so as to include not only the measurement of intelligence and achievement, but also the non-intellectual qualities and characteristics variously known as character, conduct, traits, personality, emotionality, etc. At the present time, there is intense interest in these newer developmnets and much excellent work has been done. But in some directions this movement has progressed further than it has in other directions.

In a review of the present status of the development of measuring and diagnosing personality and conduct, a number of pressing questions were revealed which are still without definite answers,[1] and there is need for much patient comparison and checking which will refine and polish methods which have been barely opened up. For instance, in the development of instruments of testing there have been noteworthy accomplishments, whereas in checking up on the reliability of the various methods and devices much remains to be done. Questions have been asked on tests and questionnaires with little consideration of the best form; tests and questionnaires have been named without a full check on the validity of the relationship between that which is measured and that which is named.

Likewise, in general, there is only dim realization of the practical possibilities in using these measures to help individuals in distress, and to understand the effects of many of our educational procedures. We are ready now to obtain accurate knowledge of the "concomitant outcomes" of education, as they have been called by Kilpatrick. Diagnostic techniques of personality and conduct should be tried out in vocational

[1] A book to be published in January, 1932, by the Century Company entitled *Diagnosing Personality and Conduct.*

selection. Studies of the extent to which emotions are aroused in various situations, both in school and out, need to be made.

In the following list of problems needing solution are many which have already been studied. They are included, not because the existing investigations are not sufficient, but because there is ample room for verification of previous studies and for new studies under different conditions and with different controls in which the influence of small variations in the situation are noted. As one instrument becomes perfected it is the signal for the repetition of an older experiment which was undertaken with cruder and less accurate tests and measures.

In reviewing the various problems which have been assembled, there appears a regularity in the type of question. The same issue arises again and again in new fields and with new materials. These may be grouped into the following fourteen categories which represent the main types of problems which demand investigation in the diagnosis of personality and conduct:

How reliable is a test, questionnaire, or other device for securing information regarding personality and conduct?

What conditions affect reliability?

How valid is a test, questionnaire, or other device for securing information regarding personality and conduct?

What is the effect of changing some elements of the situation in any one of the diagnostic techniques considered? What is the influence of various factors on test results?

What is the relation between various measures and estimates of personality, conduct, emotion, etc.?

What are the possibilities for improving the various diagnostic techniques?

What qualities are necessary for success in the investigation?

The development of new tests, questionnaires, and other devices.

What is the distribution within a group of amounts of a trait, interest, attitude, adjustment, etc.?

The problem of framing items, questions, etc.

The problem of selecting items.

The problem of weighting items.

The problem of scoring.

Differences between groups—sex, age, race, grade, environment, etc.

A List of Problems Needing Investigation in Diagnosing Personality and Conduct

OBSERVATION

1. How reliable are the various types of records, diaries, etc., for recording observation?
2. How many children can be observed simultaneously by one person without loss of reliability?
3. How does the entry of the observer into the situation as a tester or questioner affect the validity of observation?
4. Is the presence of the observer a disturbing element in observation? How serious is this factor? How long does it persist? To what extent does it depend on the observer? On the situation?
5. What is the relation of reliability of observation to the definition of the item to be observed?
6. What is the relation of reliability of observation to the number of items being observed at a given time?
7. What is the relation of reliability of observation to the subjectivity of the trait to be rated?
8. How is the reliability of observation influenced by the spacing of observations over varying periods of time?
9. What is the relation of the time interval unit of observation to the quality or behavior observed? What are the optimum time units for observing different qualities, or types of behavior?
10. What is the validity of observation schedules when prepared by different techniques?
11. How is the reliability of a directed observation as conditioned by knowledge of, or understanding of the prepared schedule?
12. What is the relation of a person's own behavior or skills to ability in carrying out a *finding* observation? Is the person whose skill is most highly developed best able to observe others?
13. What is the relation of a person's knowledge to ability in carrying out a *finding* observation? Is the person who has the most knowledge of a field best able to carry out a *finding* observation in that field?
14. How does observation of human behavior and qualities under excitement compare with observation when calm?
15. How reliably can behavior be interpreted? How accurate are interpretations from observation of human qualities, motives, and the like? How accurately does observation indicate traits or behavior trends?

16. How accurate is the memory in the observation of human qualities and behavior?
17. What is the possibility of improvement in accuracy of observation of human qualities?
18. What qualities in the individual are requisite for effective observation under various conditions?
19. Development of the techniques of observation in the classroom.

RATING METHODS

20. Is the halo effect primarily a matter of traits being rated, of the individuals doing the rating, or of the individuals being rated?
21. What is the relative influence of the acquaintance factor to other factors?
22. What is the relative influence of overestimation to other factors in self-rating?
23. What is the relation between intelligence and rating various traits in different situations?
24. When does the "altruism" factor operate in rating?
25. The less than perfect correlation between ratings and the true values of the qualities being rated lead to various regressions which result in over and underestimates. What are the various regressions that are found in connection with ratings?
26. Is there any relationship between extreme ratings in estimating ourselves and our personality adjustments?
27. What traits do we tend to overestimate in ourselves?
28. Is there more agreement in the upper end of the rating scale or in the lower end of the scale for various traits.
29. What are the characteristics of the good judge of others?
30. How does self-rating compare in reliability with the rating of others?
31. How does acquaintance affect the reliability of ratings? What are the conditions under which the "acquaintance factor" operates?
32. Will more extensive definitions or illustrations of the steps of a rating scale result in better rating?
33. Is a general trait more reliably rated than a specific trait?
34. Does the composite of the rating of several qualities into which a trait has been analyzed result in higher reliability than to rate the trait singly?
35. Which is more valid and reliable,—to rate one individual at a time, or to rate one trait at a time?

36. Does the sum of ratings of responses in several situations correlate well with the rating on the trait which these situations represent?
37. How do ratings of traits and ratings of habits compare as to validity and reliability?
38. Can raters be trained so as to improve in their ability to rate?
39. How much does a period of observation increase the reliability of ratings?

THE QUESTIONNAIRE

40. What is the form of the distribution of interests?
41. What is the form of the distribution of attitudes?
42. What is the relation between the reliability of a psychological questionnaire and intelligence of the group answering it?
43. How shall items be weighted in a psychological questionnaire? What added advantage does weighting yield over weighting all items equally?
44. What is the most advantageous scoring method when the answer to each question has been placed on a graphic rating scale?
45. How does reliability as determined by correlating two halves of a psychological questionnaire compare with the reliability as determined by a retest?
46. An analysis of the phenomenon known as "weighting the probable error" as when one validates a test and determines the scoring key on the same group in which he selects his items.
47. Is the correlation of recalled experience with present adjustment due to the true relation between these past experiences and present adjustment or because the same present forces make a person act and say things that are related?
48. What is the optimum number of alternate responses in the answers to a psychological questionnaire? Is this related to the number of items in the questionnaire?
49. How should psychological questionnaires be interpreted and scored when items have been omitted?
50. How does the true-false or yes-no question compare with questions which permit three or more degrees of judgment?
51. To what extent are the errors in a psychological questionnaire constant or variable errors?
52. To what extent does self-interest cause questionnaires to be answered falsely?

53. To what extent are individuals able to recall correctly their own past experience over carrying intervals of time?
54. To what extent are individuals able to observe their own acts?
55. To what extent is suggestion a force in the answers given in an oral interview as compared with a standardized questionnaire?
56. To what extent does the interview involve bias on the part of the examiner as compared with the standardized questionnaire?
57. What difference does a knowledge of the purpose of a psychological questionnaire make on the part of the person answering it?
58. The development of an interest questionnaire for girls, to be used for determining girl's choice of curriculum in the high school.
59. What is the value of positive statements as compared with negative statements?

THE ADJUSTMENT QUESTIONNAIRE

60. What are the relations of measures of ascendency-submission to sex, age, occupation, race, scholarship, leadership, intelligence, etc.?
61. What are the relations of measures of introversion-extroversion to sex, age, occupation, race, scholarship, intelligence, leadership, etc.?
62. What are the relations of measures of psycho-neurotic tendencies to sex, age, occupation, race, scholarship, intelligence, etc.?
63. What is the relation of introversion-extroversion questionnaires to observation of behavior as indicative of introversion-extroversion?
64. What types of maladjustment are found under varying conditions and in varying circumstances as shown by various forms of adjustment questionnaires?
65. What is the relation of scores on the psycho-neurotic questionnaire to other criteria of adjustment?
66. What is the relation between answers on a psychological questionnaire and reading ability?
67. The development of questionnaires to show adjustment in various specific situations.
68. To what extent is it valuable to ask questions on childhood and youth in the psycho-neurotic inventory?
69. How do questions on types of situation compare with questions on types of responses in the psycho-neurotic inventory?
70. To what extent is it desirable to include degrees of four or five steps in the answers to questions in the psycho-neurotic inventory?
71. What changes in the psycho-neurotic inventory can be produced by changing the examination or the school situation?

72. To what extent does the total situation,—examiner, preliminary directions, occasion (Freshman week, following an intelligence test, etc.) influence the results of a psycho-neurotic inventory?

ATTITUDE QUESTIONNAIRES

73. Experimental studies of the influence of various educational procedures,—the movies, assembly programs, lectures, posters, class discussions, etc., on attitude.
74. Experimental studies of the relationship of various social factors and influences to changes in attitudes.
75. Correlation studies of the relationship of various factors and influences to attitude.
76. What are the relations between attitude held and other measures of personality,—intelligence, information, personality and behavior ratings, interest, adjustment, emotionality, etc.?
77. What is the relationship between certainty of opinion and extremeness of attitude?
78. What is the distribution of attitude within a group?
79. To what degree does an attitude scale (Thurstone method) retain its identity when tried out on groups possessing various extreme attitudes?
80. What are the criteria for inclusion of statements in attitude scale (Thurstone method)?
81. What is the reliability of Thurstone's method of attitude measurement?
82. What is the consistency of conservativeness and radicalness over a variety of issues?
83. What is the relation between questionnaires on "How I Feel" and those on "How I Think"?
84. Construction of attitudes questionnaires on a variety of issues,—health, the family, politics, education, religion, sex, etc.
85. To what extent do attitude questionnaires agree with actual behavior?

INTEREST QUESTIONNAIRES

86. What is the relation of adjustment to interest?
87. What are the physiological correlates of interest?
88. Can involuntary movement be used as an interest measure?
89. Experiments with various types of directions and forms of response in interest questionnaires.

90. What factors are associated with interest?
91. Experimental studies of the relationship of various social and educational factors to changes in interest.
92. What sex differences in interest are found?
93. What age differences in interests are found?
94. Interest analysis (Freyd-Strong-Cowdery technique) for women's occupations.
95. What is the relation of interest to ability?
96. What is the permanency of interests,—interests in school studies, in vocational choices, in vocational preferences, in play, etc.?
97. Experimental studies of the relationship of various social factors and influences to changes in interest.

TESTS OF CONDUCT KNOWLEDGE AND ADJUSTMENT

98. What is relation of moral standards to such things as the frequency of the act, the seriousness of the act, motive, person affected by act, degree of personal sacrifice, etc.?
99. What is the relation of moral knowledge and group conduct?
100. What is the difference between the answer with personal reference and the abstract answer in tests of conduct knowledge, opinion, attitude, etc.?
101. What is the relation of ability to foresee consequences to possession of general principles of conduct in specific situations?
102. What is the relation of conduct choices to ability to foresee consequences in specific situations?
103. What is the relation of conduct choices to possession of general principles of conduct in specific situations?
104. What is the relation of Biblical knowledge to conduct?
105. What is the relation of conduct to knowledge, opinion, judgment, and attitude, determined experimentally by testing and coefficients of correlation?
106. Development of tests of opinion and judgment of conduct issues in specific situations: home, school, friends, work, sports, citizenship, etc.

PERFORMANCE TESTS

107. To what extent does it invalidate the results in a performance test of conduct to inform the pupil of what is being tested?
108. To what extent does a small change in the situation affect the results of a performance test of conduct?

109. Which is of more influence in determining a conduct response,—the person in the situation, the setting, the materials, or the motive?
110. To what extent does conduct transfer from one situation to another, determined experimentally?
111. How do various techniques for measuring honesty compare in reliability, objectivity, ease of administration, etc.?
112. What is the relation between fact scores vs. graded scores in the measurement of honesty?
113. What is the correlation of persistence with other measures of conduct and personality?
114. What is the correlation of suggestibility (with intelligence partialled out) with other measures of conduct and personality?

THE FREE ASSOCIATION METHOD

115. Development of the use of free association method in detecting cases of cheating or stealing in school.
116. What is correlation between free associations of siblings?
117. What is the optimum number of words in a free association list?
118. What factors should determine the selection of words in a free association list?
119. What factors should determine the order of words in free association list?
120. How can reaction time in free association be determined in groups?
121. What differences are there in free association when words are presented orally and visually?
122. What accuracy is needed in recording reaction time in free association?
123. What is the significance of observable behavior in free association?
124. What methods are desirable in recording free association?
125. What value is there in having an individual give train of associations which lead up to the response in free association?
126. What is the reliability of various classifications of response in free association?
127. What is the correlation of response to other factors in free association?
128. What is the relation of free association responses or reaction time to pupil-complexes,—brother, sister, parents, teacher, school, boys and girls, sports, studies, etc.?
129. What is the relation of free association response or reaction time to intelligence?

130. What is the relation of free association response or reaction time to age?
131. What is the relation of free association response or reaction time to school achievement (or failure)?
132. What is relation of free association response or reaction time to reputation for behavior by associates?
133. What is relation of free association response or reaction time to scores on the adjustment questionnaire?
134. What is the value of free association as a measure of personality integration?
135. What is the validity of complex indicators in free association?
136. What is the correlation of reaction time to other measures of emotion?
137. Development of technique of free association in detection of guilt.
138. Development of technique of free association in detection of fear.
139. What is the relation of free association response or reaction time and measures of introversion-extroversion?
140. What value has the story and story reproduction method in the detection of guilt and in the unearthing of complexes?

PHYSIOLOGICAL MEASURES OF THE EMOTIONS

141. How is the chemical composition of saliva, urine, and blood related to personality?
142. Standardization of the use of physiological measures of the emotions in diagnosing deception.
143. What is the relation of various physiological measures of emotions to free-association reaction time?
144. What is the relation of various physiological measures of the emotions to one another?
145. How do physiological measures of emotion change during oral recitation?
146. What is the relation of physiological measures of the emotions to cheating on examinations?
147. What factors are related to the psycho-galvanic reflex?
148. What is the reliability of the psycho-galvanic reflex?
149. What factors influence the reliability of the psycho-galvanic reflex?
150. What factors are related to changes in breathing?
151. What is the reliability of measures of breathing?
152. What factors influence the reliability of measures of breathing?
153. What factors are related to changes in pulse rate and volume?

154. What is the reliability of measures of pulse rate and volume?
155. What factors influence the reliability of measures of pulse rate and volume?
156. What factors are related to changes in blood-pressure?
157. What is the reliability of measures of blood-pressure?
158. What factors influence the reliability of measures of blood-pressure?
159. What are the physiological measures of emotions of those exhibiting behavior characteristics indicative of maladjustment?
160. What are the characteristics of physiological measures of emotions in school examination situations?

INTERVIEWING

161. What methods of recording interviews are most satisfactory and informing?
162. The regression effect in interviews. To what extent is it real and under what conditions?
163. What factors are related to the accuracy of testimony?
164. How does the first interview compare with the second interview in reliability, completeness, type of material disclosed, etc.?
165. What does attention to the non-verbal factors contribute to an interview?
166. How does using prepared questions in the interview compare with allowing the circumstances peculiar to the situation determine the questions?
167. What are the relative values of techniques for meeting resistance and lessening tension in the interview?
168. What is the effect of contradicting or opposing the subject in the interview as compared with agreeing with him?
169. What are the relative values of various devices in securing rapport in the interview?
170. How do the conditions of the interview,—third person present, friendships, etc., affect it?
171. What value has the preview in an interview?
172. What differences are noted between voluntary and involuntary interviews,—reliability, consistency, completeness, etc.?
173. To what extent is it possible to improve the quality of interviewing through training and practice?
174. What are the qualifications of a good interviewer?
175. In an interview, is it desirable to use questions first and give the

subject free expression later or to use the procedures in reverse order?

176. In an interview, is it better to start with a life history of the individual or with a discussion of the present irritating circumstances?
177. Does the anonymous questionnaire have any advantages over the questionnaire to which the name has been signed?
178. Which is the better adapted for obtaining disclosures as to feelings, moods, wishes, etc.,—the interview or the questionnaire?
179. Which gives the more information and more accurate information concerning childhood and youth,—the interview or the questionnaire?
180. Which will contain the greater number and more accurate personal admissions and revelations,—the interview or the questionnaire?
181. To what extent will two persons who listen to an interview make the same interpretations and draw the same conclusions?
182. To what extent are questions better understood during an interview than in questionnaires?
183. To what extent will two interviewers follow up the same leads during an interview?
184. Comparison of the interview with the questionnaire as to reliability and validity.

PSYCHOANALYSIS

185. Techniques for following up a free association.
186. Can defense mechanisms be used in adjustment questionnaires; i.e., can various common defense mechanisms be listed to be recognized by a person as representing his own or others' reactions?
187. Can escape mechanisms be used in adjustment questionnaire; i.e., can various common escape mechanisms be listed to be recognized by a person as representing his own or others' reactions?
188. Can rationalization be used in adjustment questionnaires; i.e., can various common rationalizations be listed to be recognized by a person as representing his own or others' reactions?
189. The use of rationalization in attitudes questionnaires.
190. Can transference be used in adjustment questionnaires; i.e., can various common forms of transference be listed to be recognized by a person as representing his own or others reactions?
191. A questionnaire study of regression; i.e., can various common forms of regression be listed to be recognized by a person as representing his own or others reactions?

192. Can identification be used in adjustment questionnaires; i.e., can various common forms of identification be listed to be recognized by a person as representing his own or others' reactions?
193. Can projection be used in adjustment questionnaires; i.e., can various common projections be listed to be recognized by a person as representing his own or others' reactions?
194. Can compensation be used in adjustment questionnaires; i.e., can various common compensations be listed to be recognized by a person as representing his own or other's reactions?

EXTERNAL SIGNS OF CONDUCT

195. What is the relation between the claims of graphologists and other measures of personality or conduct?
196. What is the relation between various hand and facial measurements and other measures of personality and conduct?
197. What is the relation between physique and the introversion-extroversion difference?
198. What is the relation between freedom from bodily defects and personal vigor and adjustment?
199. How is physique related to adjustment?
200. What is the relation of physique to vocational choice and to vocational success?
201. What is the relation of physique to leadership?

MEASUREMENT OF THE ENVIRONMENT

202. The development of a questionnaire designed to measure the social and character development possibilities of a school.
203. The development of a questionnaire designed to measure the social and character development possibilities of the home.
204. The development of a measure of the community environment using a technique similar to that used in the Sims Scale.
205. The development of a measure of the school environment using a technique similar to that used in the Sims Scale.

THE ARITHMETICAL BACKGROUNDS OF YOUNG CHILDREN

CLIFFORD WOODY
University of Michigan

PURPOSE OF THE INVESTIGATION

THE purpose of this report is to present some results obtained from an investigation designed to ascertain facts concerning the amount of arithmetical knowledge and skill possessed by children in the primary grades at the time at which formal instruction in arithmetic is introduced. For this investigation the time at which formal instruction is introduced was specified as the time at which a definite period in the teaching schedule is set aside for presenting facts and concepts according to a definitely schematized plan. The emphasis on systematized instruction in a period specifically set aside for such instruction differentiates it from incidental instruction which is given in response to the number needs often encountered in primary classrooms. Furthermore, formal instruction was described as planned and organized, but it was pointed out that it did not necessarily imply the use of a textbook.

TABLE I

OUTLINE OF CONTENT OF INVENTORY TEST

Type of Exercise	Number of Exercises
PART I	
Rote counting without objects	5
Counting by enumeration	2
Recognition of number as a group	
Domino arrangement	4
Dot arrangement	4
Line arrangement	4
Star arrangement	3
Reading of numbers	
Numbers on a calendar	6
Numbers in a row	6
Numbers in columns	12
Pages in a book	6
PART II	
Size of numbers	
Numbers with concrete objects	8
Numbers in the abstract	14
Telling time	
Reading time from face of clock	4
Placing hand to designate time	4
Fractions	
Recognition of fractions	3
Size of fractions	8
United States Money	
Recognition of coins	5
Value of coins	12
Linear measure	4
Liquid measure	4
Solving simple verbal problems	18
PART III	
Exercises in addition	39
Exercises in subtraction	30
Total	205

THE INVENTORY TEST

The data presented in this report were obtained through the administration of an inventory test consisting of 204 different arithmetical situations to which the child was asked to make response. Table I indicates something of the nature and number of the different types of exercises involved. In selecting the items for the test consideration was given to the following points: (1) the test should include only those types of situations which exist in the social environment of the child; (2) the test should include a great variety of situations so as to provide as complete an inventory as possible and to determine the children's familiarity with the various aspects of the number concept as outlined by Thorndike;[1] (3) the test should be simple enough for use with children in the Kindergarten and in Grade 1 and at the same time difficult enough for use with children in Grades 2 and 3. The various test items were selected only after the most careful scrutiny of the literature dealing with the teaching of primary arithmetic and of the published list of objectives enumerated for the primary grades.

THE ADMINISTRATION OF THE TEST

In general, the tests upon which this report is based were administered in May, 1929 by primary teachers in thirty-nine different school systems widely scattered throughout the United States. In one or two cities the tests were administered near the beginning of the first semester of the 1929-30 school year. In these cities the children interviewed were in the grade in which formal instruction was to be introduced. In ten of the larger cities over one hundred pupils were interviewed, but in most cities the number was considerably less than one hundred. In all but one city, no children except those in the grade just previous to that in which formal instruction in arithmetic was begun were interviewed. As a rule, each primary teacher interviewed only six pupils selected at random from those in her room. In one city, however, the following variations in procedure existed: trained specialists instead of the regular primary teachers interviewed the pupils; all pupils in Grade 1A in all class rooms instead of just six pupils in each class were interviewed; a number of pupils in all of the primary grades were interviewed in addition to those in the grades just previous to the one in which formal instruction is introduced. These variations were introduced for the sake of securing

[1] Thorndike, Edward L. *Psychology Arithmetic.*

more complete observations of the pupils' responses than untrained observers were apt to get and for securing comparative data to aid in interpreting the facts obtained in the grades just previous to those in which formal instruction is introduced.

It should be emphasized that the test was merely an interview test and in no sense standardized. For purposes of administration, it was divided into three parts: Exercises I to IV constituting Part I; Exercises V to IX, Part II; and Exercises X and XI, Part III. The test was administered individually and had no time limit, although those responsible for the interviews indicated that on the average each child took approximately twenty minutes for each part of the test, depending somewhat upon the rate at which the child responded to the different exercises and the amount of time required for the person administering the test to record observations on the pupil's behavior in making the responses. In the interview, all instructions were given orally. The child was given only one opportunity to respond to each item but he was allowed all of the time he desired in making his response. If it was evident, however, that he was unable to make a ready response he was confronted with the next exercise. During the interval of making a response the person conducting the interview recorded whether he made the proper response and made notes on any outstanding characteristics of his behavior.

RESULTS

Two different types of tabulations will be presented: tabulations showing the distributions of scores on the tests as a whole and tabulations showing responses to given sections of the test. After presentation of these tabulations consideration will be given to some observations made by those conducting the interviews.

Responses of the children to the tests as a whole.—Table II exhibits the distribution of the number of correct responses made by the children in the different grades. The same facts are presented in the form of a percentile graph in Figure 1. Before discussing the significance of the distribution of the scores, it should be pointed out that consideration of the number of pupils interviewed reveals that by far the greatest number of children were in Grade 1A, indicating that in a majority of school systems formal instruction in arithmetic is begun at the beginning of Grade 2. The next most common practice is to introduce such instruction at the beginning of the second semester of Grade 1. It is evident

TABLE II

SCORES ATTAINED IN THE DIFFERENT GRADES

Score	Kindergarten			Grade 1B			1A			2B			2A		
	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
200–209														1	1
190–199							1		1				1	1	2
180–189					1	1	2	2	4				3	3	6
170–179				1		1	10	8	18	1	1	2	11	7	18
160–169	1		1	5		5	27	14	41	1		1	17	6	23
150–159				3		3	41	26	67	3	2	5	11	17	28
140–149		1	1	5	3	8	83	44	127	5	2	7	22	19	41
130–139	1		1	9	4	13	71	56	127	1	5	6	19	14	33
120–129	1	1	2	8	7	15	81	87	168	6	3	9	13	17	30
110–119	1		1	14	9	23	98	80	178	4	5	9	12	14	26
100–109	1		1	20	18	38	83	104	187	4	5	9	6	7	13
90–99	2	1	3	19	19	38	88	113	201	3	4	7	1	8	9
80–89	2	2	4	29	21	50	93	110	203	5	4	9	3	1	4
70–79	6	7	13	41	41	82	92	94	186	4	6	10			
60–69	5	7	12	43	31	74	58	92	150	1	2	3		2	2
50–59	7	6	13	32	44	76	65	69	134		2	2			
40–49	7	11	18	40	34	74	23	29	52	1		1	1		1
30–39	9	9	18	36	30	66	13	10	23						
20–29	2	3	5	12	19	31	5	3	8						
10–19	1		1	5	2	7	2		2						
1–9				1	1	2									
Number	46	48	94	323	284	607	936	941	1877	39	41	80	120	117	237
Median	55.7	51.7	53.8	68.3	63.9	66.4	103.5	95.6	99.0	113.8	105.0	108.9	142.3	136.8	138.0
Lower Quartile	39.4	40.0	39.7	46.7	45.6	46.2	77.4	73.3	75.4	87.5	80.6	84.4	125.4	118.0	121.4
Upper Quartile	75.8	70.0	72.7	91.8	85.2	88.6	130.1	120.2	123.8	140.5	127.5	131.7	151.2	153.4	156.7
Significance of Difference of Means		1.6			0.4			1.3			1.8			—0.6	

from this tabulation and from the original interview blanks that in some schools formal instruction is begun during the first semester of Grade 1; in other systems it is postponed until the beginning of Grade 3. Such observations are in keeping with other existing facts concerning the time for introducing such instruction.

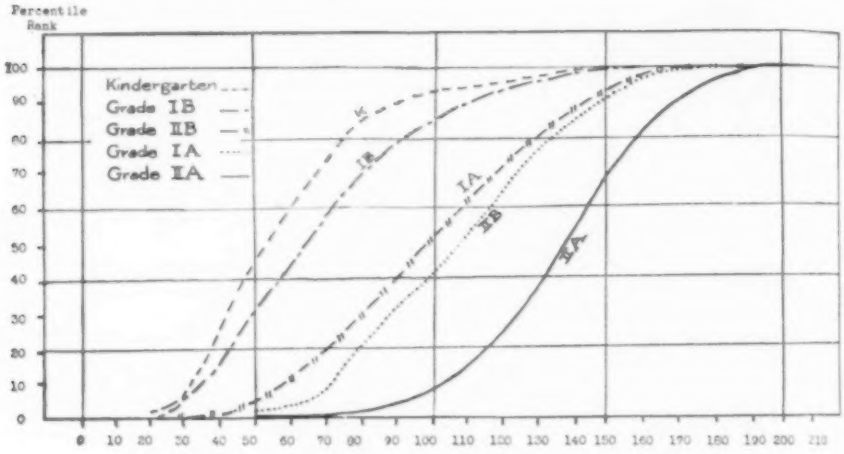


FIG. 1. SCORES ATTAINED IN THE DIFFERENT GRADES

Probably the most outstanding point portrayed by Table II and its graphic presentation is the wide range in the number of correct responses given by the children who responded correctly to a very small number of the items of the test and by those who responded to a surprisingly large number of the exercises. It is reasonable to expect some range in the achievements of the children in a given grade, but it is rather surprising to find that the ranges in the Kindergarten and Grade 1B are almost as large as in Grades 1A, 2B, and 2A. Some overlapping of the responses of the pupils in the various grades was to be expected, but the number of correct responses made by the children in the Kindergarten and Grade 1B was not anticipated. Incidentally, it may be added that teacher after teacher in one of the school systems commented to those administering the tests that it was useless to interview the children in the given grade because so many of the test items were beyond the abilities of their children.

Examination of the median scores indicates that the children in the Kindergarten responded correctly to approximately one-fourth of the exercises; the children in Grade 1A, to nearly one-half of them; and

the children in Grade 2A, to about three-fourths of them. These facts are all the more significant when it is recalled that the children under consideration had received no formal instruction in the school. The progress between Grades 1B and 1A and between Grades 1A and 2A is rather outstanding as indicated by the wide differences between their respective medians and the percentile curves showing the achievement of the respective grades. There is not so much difference between the achievement of the children in the Kindergarten and Grade 1B and between that for the children in Grades 1A and 2B. This may be due to the unreliability of the data for the children in these grades or to the fact that the children in the B-sections of these grades, i.e., those entering the respective grades at mid-year, make relatively lower scores than the children entering the grades at the beginning of the regular year.

The data indicate that the median number of correct responses for boys tends to be higher than the median number of correct responses for the girls. The difference in the number of correct responses, usually six or eight exercises, is not statistically significant as indicated by the coefficients of the significance of the difference given at the bottom of the table, yet in all grades the number of correct responses is greater for the boys than for the girls. In tabulations showing the percentages of boys and girls responding correctly to each group of exercises within the test, the percentage for boys is again higher than that for girls in a large majority of the types of exercises. This fact would lead one to believe that this difference in the responses of the boys and girls may have more significance than the coefficients suggest.

Table III is similar to Table II, but it contains only the responses made by six-year old children in the various grades. A six-year old child was defined as one whose age was between five years, six months, and six years, six months. The facts of this tabulation are shown in Figure 2. This tabulation and graphic presentation reveal about the same tendencies as presented in Table II and Figure 1, with the exception that the difference between the number of correct responses by the boys and girls in the different grades are somewhat different when the factor of age is kept constant. The boys on the whole tend to make the greater number of correct responses, but the differences are not as great as in the previous exhibits.

Responses to the exercises involving counting.—Table IV presents the extent to which the children in the Kindergarten to Grade 2A, inclusive, could count. Approximately one-fourth of the children in the Kinder-

TABLE III

Scores Attained by Boys and Girls Six Years of Age

Score	Grade: Kindergarten			1B			1A			2B			Total		
	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
200-209															
190-199															
180-189								1	1					1	1
170-179				1		1	1	2	3				2	2	4
160-169	1		1	3		3	6	7	13				10	7	17
150-159				3		3	12	10	22				15	10	25
140-149				3	2	5	23	25	48		1	1	26	28	54
130-139	1		1	7	3	10	18	15	33				26	18	44
120-129		1	1	6	5	11	25	36	61				31	42	73
110-119	1		1	12	7	19	45	24	69				58	31	89
100-109	1		1	12	12	24	34	46	80	1		1	48	58	106
90-99		1	1	13	13	26	33	40	73				46	54	100
80-89	2	1	3	18	19	37	27	38	65				47	58	105
70-79	4	4	8	22	30	52	28	48	76	1	1	2	55	83	138
60-69	4	3	7	25	21	46	16	29	45				45	53	98
50-59	5	3	8	20	27	47	20	23	43				45	53	98
40-49	5	7	12	22	22	44	9	3	12	1		1	37	32	69
30-39	4	2	6	26	19	45	5	2	7				35	23	58
20-29	1		1	10	10	20	1	1	2				12	11	23
10-19	1		1	2	1	3							3	1	4
1-9					1	1								1	1
Number	30	22	52	205	192	397	303	350	653	3	2	5	541	566	1107
Median	57.8	56.7	57.5	69.0	67.6	68.4	103.7	97.8	100.4	75.0	110.0	77.5	83.2	84.5	86.1
Lower Quartile	43.0	45.0	44.2	46.0	47.7	46.9	78.8	76.1	77.1			71.2	60.7	63.9	62.4
Upper Quartile	76.9	73.8	75.0	96.7	86.8	91.1	123.7	122.4	122.9			107.5	115.6	110.0	113.4
Significance of Difference of Means		0.8			1.2			0.5			—			0.3	

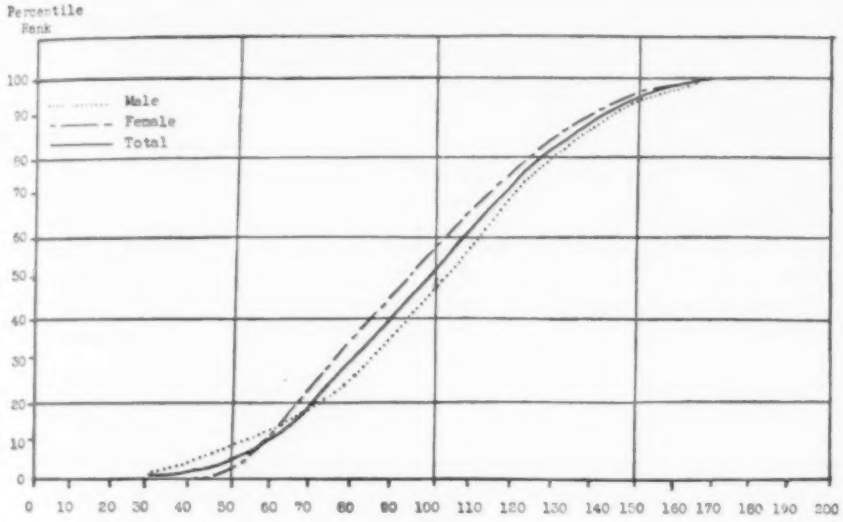


FIG. 2. SCORES ATTAINED BY BOYS AND GIRLS SIX YEARS OF AGE

TABLE IV

PERCENTAGE OF CORRECT RESPONSES GIVEN TO EACH COUNTING EXERCISE BY THE CHILDREN IN THE PRIMARY GRADES

Item	Grade				
	Kindergarten	1B	1A	2B	2A
1. Rote counting to 100 by 1's	26	38	66	76	94
2. Rote counting to 100 by 10's	32	46	69	76	93
3. Rote counting to 30 by 2's	7	14	34	28	68
4. Rote counting to 30 by 3's	3	2	9	11	16
5. Rote counting backward from 20 by 2's	4	3	12	12	35
6. Counting 20 circles and pointing order while counting	71	79	93	98	97
Number of children	94	604	1897	80	238

garten, almost two-fifths in Grade 1B, two-thirds in Grade 1A, three-fourths in Grade 2B and over nine-tenths in Grade 2A counted to 100 by 1's. About the same proportion of the children counted correctly to 100 by 10's. Approximately three-fourths of the children in the Kindergarten and Grade 1B counted the twenty circles and pointed at them in proper order as they counted. In Grades 1A, 2B, and 2A over nine-tenths of the children responded correctly to this exercise. Some children in each grade responded correctly to each of the other exercises, but the percentage of correct responses while increasing from grade to grade were much smaller than those for the enumerated exercises.

To the writer the data presented in this table reveals four outstanding facts: (1) the children have considerable knowledge of counting even though formal instruction in the subject has not begun; (2) the percentage of correct responses increases from grade to grade; (3) the exercises involving counting twenty circles and counting them in order proved easier than rote counting to 100 by 1's or 10's; (4) the exercise involving rote counting to 100 by 10's was easier than that involving counting to 100 by 1's.

Before passing to the next group of exercises it may be well to point out that on the set of six exercises and on each of the items in almost every grade there was a slight superiority in the percentage of correct responses in favor of the boys.

Responses to exercises involving the telling of time.—The first four exercises involving time required merely the reading of the time as indicated on drawings of the face of various clocks; the last four exercises, the placing of the minute hand on the clocks to indicate designated times of day. The most significant facts portrayed in Table V are: (1) the

TABLE V

PERCENTAGE OF CORRECT RESPONSES TO EACH EXERCISE INVOLVING THE TELLING OF TIME MADE BY THE BOYS AND GIRLS IN THE PRIMARY GRADES

Items	Kindergarten		Grade 1B		1A		2B		2A	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1. Clock says 9 o'clock	41	40	32	29	53	49	55	50	72	78
2. Clock says 1 o'clock	52	44	37	34	56	51	52	45	74	70
3. Clock says 12 o'clock	41	38	49	44	69	67	64	63	91	96
4. Clock says 3:30 o'clock	7	4	8	5	19	12	17	13	44	41
5. Place hand at 7 o'clock	22	19	26	21	47	42	36	42	75	78
6. Place hand at 5 o'clock	24	21	22	19	44	39	31	42	73	75
7. Place hand at 10 o'clock	20	19	24	19	44	41	33	39	74	80
8. Place hand at 11:45 o'clock	7	2	11	4	14	10	19	8	38	36
Number of pupils	46	48	315	289	929	968	42	38	117	121
Mean percent	27	23	26	22	43	39	38	38	68	69

mean percentage of correct responses on all of the exercises for the children in the Kindergarten and Grade 1B was approximately 25 percent; for the children in Grades 1A and 2B, slightly over 39 percent; for Grade 2A, 69 percent; (2) the exercises involving the reading of the time were considerably easier than those in which the placing of the minute hand to indicate a given time was required; (3) on most of the exercises, especially the more difficult ones, the boys had a higher percentage of correct responses than the girls.

Responses to exercises involving fractions.—From Table VI there is evidence that children in the primary grades have considerable familiarity with the fractions: one-half, one-third, and one-fourth. Approximately two-thirds of the children in the Kindergarten and Grade 1B, three-fourths in Grades 1A and 2B, and nine-tenths in Grade 2A selected from the set of pictures of apples cut into various fractional parts, the one which showed the apple cut into halves. Almost half of the children in the Kindergarten and in Grade 1A, and slightly over two-thirds of the children in the other grades under consideration were able to select the apples cut into fourths and thirds. The percentages of children re-

TABLE VI

PERCENTAGE OF CORRECT RESPONSES TO EACH EXERCISE INVOLVING FRACTIONS MADE BY THE BOYS AND GIRLS IN THE PRIMARY GRADES

Items	Kindergarten		Grade 1B		1A		2B		2A	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
POINT TO APPLE										
Cut into halves	63	65	68	65	77	77	86	76	89	86
Cut into fourths	41	48	52	52	67	68	67	71	74	74
Cut into thirds	37	42	45	46	67	63	60	63	75	78
YOU CAN CUT AN APPLE INTO										
How many halves	50	46	45	42	44	51	55	29	50	41
How many fourths	30	17	30	27	35	36	52	29	48	44
How many thirds	26	17	24	20	33	35	33	26	50	44
WHICH IS LARGER?										
Half or whole apple	80	88	85	85	94	93	100	100	97	98
Fourth or half apple	39	42	43	42	49	46	69	47	71	59
Half or third of apple	17	19	35	29	46	43	52	61	74	69
Fourth or third of apple	41	35	35	36	33	33	29	39	25	28
Half, third, or fourth of apple	15	10	20	18	35	34	52	32	62	63
Number of pupils	46	48	315	279	929	968	42	38	117	121
Mean percent	40	39	44	41	53	53	60	52	65	62

sponding correctly to the questions involving knowledge of the number of halves, thirds, or fourths into which an apple can be cut was considerably less than the percentage responding correctly to the exercises involving the designation of apples cut into enumerated fractional parts. The responses to exercises involving the relative size of the fractions indicate that a very large proportion of the children knew that a whole apple is larger than half an apple; approximately half of them knew that half of an apple is larger than a fourth of an apple. A considerably smaller proportion of the children responded correctly to the exercises involving half and third, fourth and third, and half, third, and fourth, but the percentage making correct responses indicates some understanding of the meaning of size of these fractions. The responses to the last five of the exercises involving the size of fractions are somewhat subject to error arising from the form of the test exercises, but those responsible

TABLE VII

PERCENTAGE OF CORRECT RESPONSES TO THE EXERCISES IN ADDITION MADE BY THE CHILDREN IN THE PRIMARY GRADES

Exercise 10	Grade: Kindergarten		1B		1A		2B		2A	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
2+1	43	38	57	46	81	76	88	87	92	98
0+2	30	31	40	30	61	51	74	63	80	84
6+2	26	13	32	22	61	53	74	63	87	87
7+1	33	19	42	30	68	63	81	68	85	93
2+2	41	27	47	37	74	70	90	74	92	99
7+1	33	17	41	29	69	62	83	68	91	95
6+3	15	13	28	20	57	49	69	58	85	84
5+4	15	8	22	14	53	42	71	58	84	83
5+7	4	4	12	6	34	26	45	42	63	64
6+4	4	6	21	10	47	39	50	50	76	72
3+9	7	8	13	6	36	28	33	45	66	69
9+1	13	10	30	20	58	52	67	61	88	90
5+8	6	4	13	3	30	23	36	37	65	55
5+5	13	8	30	19	58	51	60	63	86	90
9+8		2	7	3	18	14	24	18	56	40
7+6	4	2	10	3	28	20	36	32	62	52
2+3+3	9	8	16	8	41	32	38	42	72	67
4+1+2	11	4	17	8	43	35	50	42	75	72
2+3+1	11	4	14	8	40	31	40	45	76	72
4+3+2	11	2	16	7	38	30	45	37	71	64
5+1+3	4	6	16	6	40	31	50	45	74	70
2+4+9		2	7	3	19	15	33	29	57	47
4+5+9			5	2	17	12	19	13	50	37
2+5+8			5	2	18	13	26	21	57	44
12+3	11	6	12	6	33	26	36	42	68	67
14+4	4	2	9	4	26	23	17	37	65	60
21+7	2		3	1	11	8	12	11	39	32
17+2	4	4	11	4	29	24	26	34	71	64
12+8		4	5	1	16	13	17	24	44	40
16+5	4	2	7	2	20	15	24	18	58	51
29+4	2	4	6	1	16	14	14	29	50	46
38+9			3	0	9	8	5	16	36	33
22+13+31			2	0	3	4		5	23	24
41+15+23			2		3	4		8	18	25
54+33+12			1		3	3		5	19	25
57+84+79			0						6	2
324+453			1		3	2			17	21
576+425+358								5	3	6
987+456+321+237									3	2
Number of pupils	46	48	315	289	929	968	42	38	117	121
Mean percent	9	7	15	9	32	27	37	36	59	57

for the interviews were instructed to make a record of the pupils who seemed to be guessing and who had no understanding of the responses made. In general, the facts in this table suggest that children in the primary grades have some knowledge and understanding of such fractions as halves, thirds, and fourths, and that such concepts are within the comprehension of such children. Incidentally, in almost all grades, it is interesting to note that on this set of exercises and on most of the individual items the boys gave a slightly higher percentage of correct responses than the girls.

Responses of the children to the exercises involving addition.—Table VII indicates a great range in the knowledge of addition possessed by children in the primary grades before formal instruction was introduced. Consideration of the mean percentages on all of the items or on any one of them shows that the amount of knowledge increased from grade to grade. The children in the kindergarten manifested considerable knowledge of the easier combinations, some knowledge of the harder ones, and a little knowledge of the exercises involving adding three one-place numbers and those involving a two-place number and a single place number. In each of the successive grades an increasingly greater percentage of the pupils gave the correct responses to these types of exercises and to those types representing more difficult combinations and more complicated processes. The children in Grades 1A, 2B, and 2A ready for the introduction of the formal study of arithmetic in the school systems under consideration at the beginning of the next school year, possess, if their scores mean anything, a surprising amount of knowledge of the subject matter to which they will be exposed during their first year of systematized instruction. Other tabulations not presented in this discussion but exhibiting the distribution of responses made by individual children on the sets of exercises, reveal that some children in each of the grades have given almost one hundred percent responses to all the types of exercises which will be taught during the first year of formal instruction. The variation in the percentages of correct responses to each of the items is enlightening, but not quite so much as is the variation made in the responses by individual pupils.

Space will not permit a more detailed analysis of the facts given in Table VII, but attention should again be called to the fact that the boys in each of the grades under consideration made a greater percentage of correct responses on the test as a whole and on most of the items in the test.

OBSERVATIONS MADE DURING RESPONSES BY THE CHILDREN

The facts presented in the previous sections represent the final responses made by the children, but it should be pointed out that these responses should not be taken at one hundred percent face value. In many cases the record finally given a child was indicative of his real achievement, but in other cases it was a poor indication. Some children responded accurately with rapidity and assurance; others responded slowly with frequent self-corrections and by means of "crutches."

A few illustrations will show the necessity of caution in interpreting the results obtained. In rote counting, some children after completing the exercise offered to count most satisfactorily to 100 by 5's, to count above 100, or to count backwards from thirty by 1's; others, while finally stumbling through with no more aid than general encouragement, counted rapidly to 20 or 30 or even 40 and then paused, repeated the count of 40 two or three times, started again counting rhythmically for a time, and then repeated the faltering process. Some children responded correctly in rote counting to 100 by 10's and to 30 by 2's or 3's, but at the same time repeated the intervening numbers in an audible whisper. In fractions, some children responded correctly to the questions concerning the size of halves and fourths, but had to look at the pictures of the apples cut into fractional parts before answering; one child stated that a whole apple was larger than a half apple unless the apple was a very small one; still another stated that an apple could be cut into halves, but added that if the apple were very large it could be cut into three halves. In addition, many children gave the correct response to the different exercises but determined the response through counting. Some counted aloud, others silently; some counted their fingers, objects in the room, cherries in the pattern of dress, or by the making of marks on paper. A few did double counting, e.g., in adding 8 and 5 they would count by rote to 8 and then say 8 and 1 more are 9, 8 and 2 more are 10, etc., until they had counted off 8 and 5. Other illustrations might be cited, but enough have been given to suggest that the interpretation of the final responses made to exercises like those constituting this interview blank must be conditioned not only by themselves but also by the method employed in arriving at the responses.

CONCLUSIONS

From this study five outstanding conclusions are warranted: 1. that children possess much ability in the elementary processes of arithmetic

even before the time of beginning formal instruction in the subject; 2. that the knowledge possessed by children is not limited to counting and adding simple combinations, but includes elementary knowledge of fractions, United States money, units of various types of measurement, and the understanding of the processes demanded in simple verbal problems; 3. that extended inventory tests designed not only for throwing light on the goals of achievement, but also upon the methods employed in attaining the goals should be given before the beginning of formal instruction if instruction is to be adapted to the needs of the individual child; 4. that the method of attaining a solution must be taken into consideration in interpreting the significance of a given attainment; and 5. that a small but persistent tendency exists for boys to make a slightly greater percentage of correct responses to tests as a whole and to individual items of the test than girls.

A METHOD FOR MEASURING RETENTION IN READING

ALVIN C. EURICH
University of Minnesota

READING has no particular merit in broadening the experiences of an individual unless what is read is retained. One may comprehend while reading and yet, immediately following the process, not remember all the ideas that were symbolized by the printed words. This phenomenon practically forces one to believe that comprehension and retention are two separate functions of the reading process. The literature is confusing at this point, since a number of studies are reported in which "comprehension" is supposedly measured by the amount of a passage an individual can reproduce immediately after he has read it. Because of this unwarranted and unrestricted use of the term "comprehension" there is need for an attempt to isolate certain factors that have been obscured by such loose terminology. The method involved in the Minnesota Reading Examination Xi XIII[1] is one effort in this direction, since the test is concerned with the measurement of the information which a student retains from the reading of a given passage.

The Construction of the Test.—Reading passages that were of a fair degree of difficulty for college students and that contained a proportionately large number of facts or items of information were sought. After a perusal of a number of books which had been written for the college population, it was decided to make the selection from Newman and others, *The Nature of the World and of Man.* To provide the reading material for Form A of the test, pages 191 to 197 were printed in a separate booklet. These pages are concerned with a discussion of *The Evolution of Life* and are so labeled on the printed booklet. For Form B a discussion was reprinted from *The Eugenic Program* as it appears on pages 434 to 439 of the same book.[2]

The next step towards the completion of these two forms was the framing of an objective examination on the information contained in each of the booklets. The type that seemed most applicable in this instance was the multiple choice. This precluded the measurement of retention by recall and practically assured its measurement by recognition which probably indicates maximum retention rather than minimum.

[1] For more complete data regarding this test see: Eurich, A. C. *Reading Abilities of College Students.* Minneapolis: University of Minnesota Press, 1931.

[2] Permission to reprint these pages was obtained from the University of Chicago Press.

When finally formulated, the two examinations consisted of fifty units each which covered thoroughly the text of the reading passages. These examinations were printed in separate folders to permit their use under varying conditions.

Methods of Scoring.—For scoring purposes, each of the fifty units in the test was defined as a task and one point was given for each. For some tasks, successful completion required that two or more phrases be underlined, whereas in other tasks one phrase was sufficient. Because of this fact the method of scoring that was adopted may be open to question. In order to note its agreement with other methods of scoring, one set of papers was scored in three ways. First, the number of correct units was determined. Second, one point was given for every item correctly underlined regardless of the units. In other words, if one unit contained two items it was given two points. Third, one point was assigned for each item correctly underlined and one point was subtracted for each item which was wrong. As a measure of the agreement between the first method of scoring and the other two, the coefficients of correlation were obtained. That found between the first and second methods is .89 ± .02. Between the first and third methods it is also .89 ± .02. These values suggest fairly close agreement; in fact, the coefficients are higher than the reliability of the test itself. Since the relation is as close as indicated, the first method was adopted for future scoring as it took somewhat less time and was simpler to persons scoring the tests.

Reliability.—With this scheme for scoring, the reliability values were first sought in the correlations between the two forms. Since the tests were given before the students read the passages and again immediately after they had read them, the correlations as presented in Table I have due regard for these two periods of testing. The coefficient of correlation

TABLE I

EXAMINATIONS XI XIII—RELATION BETWEEN THE EXAMINATIONS ON THE EVOLUTION OF LIFE (A) AND THE EUGENIC PROGRAM (B)

Time of Testing	r P.E.
Before Reading	.34 ± .05
After Reading	.53 ± .04

between Forms A and B before reading is .34 ± .05. After reading this becomes .53 ± .04.

One may be in error in interpreting these coefficients strictly as measures of reliability, inasmuch as the nature of the reading material

varies from one form to the next. In Form A, or "The Evolution of Life," the reading matter is rather general, whereas in Form B, or "The Eugenic Program," it is very specific and precise with many figures that add to its exactness. Because of this fact, the two forms may actually be measuring different factors in the reading process. Before such a conclusion can be drawn, however, further experimentation is necessary with this type of test.

Other reliability coefficients were found by securing the correlations between the two halves of the tests and projecting these values with the Spearman-Brown formula. As calculated in this manner, they appear in Table II for the test given "before reading," "after reading," and "one week later." Such an arrangement was possible, since for one group

TABLE II

RELIABILITY OF EXAMINATION Xi XIII

Time of Testing	A	B
Before Reading	.79 ± .03	.29 ± .10
After Reading	.69 ± .05	.63 ± .05
One Week Later	.84 ± .03	.74 ± .05

of college juniors the test was given at the three periods indicated. These correlations are somewhat higher than those reported between the two forms of the test. They reveal further evidence that there is a difference between the two forms. For Form A the reliability coefficient "before reading" is .79 ± .03, whereas on the same period of testing for Form B the coefficient is only .29 ± .10, which indicates that scores on the latter form are largely due to chance. Those on the first form depend upon the information concerning *The Evolution of Life* which was previously acquired. The reliability coefficients for the "after reading" scores are approximately the same magnitude for both forms of the test (.69 for Form A and .63 for Form B). The scores "one week later" are even more reliable (Form A, .84; Form B, .74). Obviously the test is least reliable before reading, more reliable after reading, and most reliable one week after reading the passage. This seems to be particularly true of the test that covers the specific material. In order to secure a better appreciation of these reliability coefficients there is need for further work with retention for specific and general material.

Validity.—As with most other tests, the validity of the Minnesota Reading Examination Xi XIII must rest almost exclusively upon its inherent nature. In all probability it approaches an actual study situation more closely than any tests which are now available. Before reading the passage the student secures an idea of its content through the ques-

tions in the test. During the reading process, he has the opportunity to collect facts with reference to that idea with the definite intention of taking an examination upon those facts as soon as he has completed his reading.

Relation to Other Variables.—The extent to which scores on this test are predictive of achievement in a limited number of college courses may be observed from the coefficients of correlation in Table III. In the

TABLE III

THE RELATION OF EXAMINATION Xi XIII TO SCHOLARSHIP

Course	Form A	Form B
Educational Psychology	.42 ± .06	.34 ± .06
How to Study	.44 ± .06	.33 ± .07
Honor Point Ratio for all Courses	.35 ± .05	

calculation of these coefficients the percent of possible gain that was made on the "after reading" test over the "before reading" test was used as the score for Examination Xi XIII. To establish the possible gain the total number of units in the test was used as the upper limit. This was 50 points. An illustration of the application of this method will aid in clarifying any obscure points in this brief description. If a student secured a score of 20 "before reading," his possible gain is 30 points. If his score is 35 "after reading," he has gained one-half of his possible gain, or fifty percent.

Certain defects of this method are at once apparent, since it is not known that anyone can secure the total number of points in the test after he has read the passage. The truth of this may make the statement of the possible gain a fallacious one. It does, however, afford an arbitrary point of reference that is constant for all.

Figures which are expressive of the degree of relationship between this test and intelligence, vocabulary, and reading comprehension appear in Table IV. If the Minnesota College Ability Test is used as a measure

TABLE IV

THE RELATION OF EXAMINATION Xi XIII TO OTHER VARIABLES

Variable	Form A	Form B
College Ability Test	.41 ± .07	.40 ± .07
Minn. Reading Examination, Part I	.21 ± .06	.14 ± .06
Minn. Reading Examination, Part II	.37 ± .05	.26 ± .06
Minn. Speed of Reading Test, Form A	.45 ± .04	

of intelligence, the coefficients of correlation of .41 for Form A and .40 for Form B denote the relationship between intelligence and the percent of possible gain on Examination Xi XIII. Since Part I of the Minnesota

Reading Examination is a vocabulary test, the coefficients .21 and .14 are indicative of the relationship between Examination Xi XIII and vocabulary. For Part II, or reading comprehension, the coefficients are somewhat larger, while that secured for the Minnesota Speed of Reading Test is the largest of all. None of these coefficients are high. Corrected for attenuation they would all be considerably higher. Whether or not they are indicative of the true relationship between these variables remains for further demonstration.

Application of the Test.—Aside from the application of this test under the experimental conditions for which the instrument was designed, it was used with several groups under varied circumstances in order to compare the relative gains. With one group of forty students, all of whom were college juniors, the initial test was given under conditions which gave each individual all the time he required to respond to the fifty units. The reading pamphlet was then given with directions to read it once. Immediately upon the completion of this reading the students turned in their booklets and the reading time was recorded. As soon as all booklets were gathered, the second test was administered under precisely the same directions as the initial test. One week following this reading the same test was given again for the purpose of determining the loss during the interim. The facts as derived are arranged in Table V. The means and

TABLE V

AMOUNT OF RETENTION AS MEASURED BY EXAMINATION XI XIII WITH UNLIMITED TIME

	Form A *The Evolution of Life*			Form B *The Eugenic Program*		
	M	P.E.	S.D.	M.	.P.E	S.D.
Before Reading	16.33 ± .45		4.2	16.58 ± .45		4.2
After Reading	28.74 ± .45		4.2	35.00 ± .50		4.6
Gain	12.41 ± .64			18.42 ± .68		
One Week Later	26.44 ± .52		4.8	32.05 ± .52		4.8
Loss	2.30 ± .69			2.95 ± .73		
Gain over Initial Test	10.11 ± .69			15.47 ± .69		

standard deviations as expressed here are in terms of raw scores. On Form A the gain in means from the initial test to the test directly following the reading is 12.41 raw score points or approximately 35 percent of the possible gain calculated from the maximum numbers of points in the test. One week later there is a drop in the means of 2.30 points. When this loss is interpreted in terms of its probable error, a ratio of 3.33 is obtained, which suggests a significant difference. The gain in the final test over the initial test is 10.11 points, which is highly significant and shows without question that the amount of retention one week after

reading a six-page article is considerable. The facts for Form B are similar to those for Form A with the exception that the mean gain is greater, being approximately 53 percent of the total possible gain. The loss after an interval of one week is also greater, as the ratio to its probable error is 4.04.

The intercorrelations that are grouped in Table VI were derived from these same data. With the exception of the coefficients for the actual

TABLE VI

INTERCORRELATIONS FOR EXAMINATION Xi XIII

Combination	A	B
Before Reading and After Reading	.37 ± .09	.45 ± .09
Before Reading and Actual Gain	—.48 ± .08	—.41 ± .09
Before Reading and Percent Possible Gain	.22 ± .10	.18 ± .10
After Reading and Actual Gain	.55 ± .07	.59 ± .07
After Reading and Percent Possible Gain	.72 ± .05	.85 ± .03
After Reading and Reading Time	—.16 ± .10	—.34 ± .10
Actual Gain and Reading Time	—.13 ± .10	—.69 ± .05
Actual Gain and Percent Possible Gain	.95 ± .01	.73 ± .05
Before Reading and One Week Later	.46 ± .08	.24 ± .10
After Reading and One Week Later	.66 ± .06	.56 ± .07
Actual Gain and One Week Later	.22 ± .10	.17 ± .10
Percent Possible Gain and One Week Later	.38 ± .09	.47 ± .08

gain as one variable and the percent of possible gain as the other, they are not high and indicate considerable shifting in the relative ranking of individuals "before reading," "after reading," and "one week later." The negative correlations between the "before reading" scores and actual gain are indicative of the fact that the students with the lowest scores on the first test made the greatest actual gains, while those with the highest scores in the first test made the smallest actual gains. The coefficients between the "after reading" scores and percent of possible gain are sufficiently high to suggest that the scores made "after reading" may be used as retention scores when it is impossible to give the examination before reading and after reading. The discrepancy between the coefficients for actual gain and reading time may also be due to the specificity of the content of Form B and the general character of Form A. The correlation for Form B clearly points to the tendency for the slowest readers to make the greatest actual gains and for the most rapid readers to make the

TABLE VII

AMOUNT OF RETENTION AS MEASURED BY EXAMINATION Xi XIII WITH LIMITED TIME

	The Evolution of Life			*The Eugenic Program*		
	M.	P.E.	S.D.	M.	P.E.	S.D.
Before Reading	10.70	± .25	4.00	12.82	± .26	3.90
After Reading	27.62	± .37	6.04	35.58	± .39	6.45
Gain	16.92	± .44		22.76	± .47	

least actual gains. Undoubtedly the nature of the reading material determines the magnitude of these coefficients to a considerable degree.

Another group of 125 college juniors and seniors were given the Minnesota Reading Examination Xi XIII with definite time limits. That is, they were given fifteen minutes for the test "before reading," fifteen minutes for reading the booklet, and fifteen minutes for the test "after reading." The data for this group as arranged in Table VII do not differ significantly from those just presented. The chief variation is the lower measure of central tendency on the initial test which in turn provides a greater gain, as these students on the average attain the same level after reading as did those individuals who were given unlimited time. Since the test was not repeated with this group one week after reading, further comparisons are not afforded.

SUMMARY

In the foregoing pages, the writer has been primarily concerned with the presentation of a method for measuring retention in reading. The proposed method forms the basis for the Minnesota Reading Examination Xi XIII. The test itself consists merely of an objective examination on the text of a six-page booklet. This examination is given before reading and after reading the passage. Two forms of the examination were constructed, one containing general material and the other specific material. When applied to several groups the results from both forms indicated considerable gain in information accruing from the reading process with a relatively small amount of loss during an interval of one week.

IMPROVING COLLEGE FRESHMEN IN SPELLING

WALTER S. GUILER
Miami University

IN order to overcome the background deficiencies of college freshmen, the School of Education of Miami University has resorted to a program of diagnostic testing and remedial instruction. This article reveals what has been accomplished in the field of spelling.

Discovering the weak spellers constituted the first step in the program. For this purpose the writer's preliminary spelling test was used. The fifty words comprising the test appear in Table I. They were selected from a basic list of 2936 words in most common use and scaled on the basis of difficulty. In the determination of the content of the basic list, only those words which were found both in Dr. Ashbaugh's *Iowa Spelling Scales* and in Dr. Horn's *A Basic Writing Vocabulary: 10,000 Words Most Frequently Used in Writing* were included. The score which a student makes in the preliminary test should thus be reliably indicative of his ability to spell the three thousand words in most common use.

TABLE I

WORDS USED IN THE PRELIMINARY SPELLING TEST

buy	advice	advanced	advisable	lieutenant
color	lining	calendar	interfere	phosphorus
place	losing	choosing	privilege	ingredients
prior	ninety	counting	principal	anniversary
ample	believe	grateful	sincerely	accommodate
apply	capital	superior	performed	appendicitis
affair	certain	mortgage	permanent	considerably
waste	circuit	employee	materially	neighborhood
except	receipt	beginning	assessment	administrator
amount	attempt	convinced	excitement	affectionately

The test was made self-ministering by embodying each word in a meaningful sentence and enclosing the word in parenthesis. Some of the enclosed words were spelled correctly; others were spelled incorrectly in their most frequently misspelled form. In taking the test, the student was instructed to look at each sentence only long enough to discover the word to be spelled, and then to write its correct spelling on the dotted line at the end of the sentence, as in the samples.

Samples: 1. John is a (tall) man. ....tall....
2. Ray is (fore) years old. ... four....

The results obtained from giving the preliminary test to 350 entering freshmen during the first week of the 1928-1929 college year are presented in Table II. The highest possible score is fifty. All of the students who did not attain a score of 44 were selected as "hospital cases."

TABLE II

SCORES MADE BY 350 COLLEGE ENTRANTS IN THE PRELIMINARY SPELLING TEST

No. of words correctly spelled	Frequency	Grade Standards: Educational level	Grade Standards: Median score
50	5		
47-49	53	Teachers	
44-46	91	in service	46.3
41-43	57		
38-40	47	College	
35-37	25	freshmen	42.8
32-34	22		
29-31	21	Grade 12	38.9
26-28	14	Grade 11	37.6
23-25	6	Grade 10	34.6
20-22	4	Grade 9	27.7
17-19	3	Grade 8	22.9
14-16	2	Grade 7	18.2
	350		
Average score	40.2*	Median score	42.7
90 percentile score	48.0	25 percentile score	36.8
75 percentile score	46.1	10 percentile score	29.9

* All computations were derived from ungrouped data.

Analysis of the undistributed scores from which Table II was constructed revealed several significant facts. First, many of the college entrants manifested superior ability in spelling; in fact, approximately one-sixth of them attained a score which is higher than the median score for teachers in service. Second, a large number of the freshmen exhibited marked deficiency in spelling ability. Out of the total group of 350 students, 31.4 percent fell below the standard for high school seniors; 6.3 percent did not attain the standard for the ninth grade, which is the last year of the junior high school; and 1 percent did not reach the standard for the seventh grade, which marks the beginning of the junior high school. Third, the students manifested a wide range in ability to spell. Five freshmen made perfect scores, while two students misspelled 34 of the 50 words in the preliminary test.

Helping each weak speller identify the particular words in the basic list which he had not mastered comprised the second step in the program of remediation. This phase of the work was based on the principle that any economical and effective plan for overcoming spelling deficiencies must first of all identify for each individual the words which he cannot spell in a list of words which he should be able to spell. Accordingly, each remedial student was given an individual diagnostic test covering all of the words in the basic list to which attention was called in the second paragraph of this article. This test comprises Part II of *A Mastery Work-Book in Spelling* and is so organized that it is in reality a

diagnostic self-test. The test was scored by the student with the help of a mastery word-list, which comprises Part I of the student's work-book. Each word missed in Part II was marked in Part I by means of a cross (X). In this way, the words in the basic list which proved difficult

TABLE III

POINTS OF DIFFICULTY ENCOUNTERED BY 350 COLLEGE FRESHMEN IN SPELLING THE WORDS IN THE PRELIMINARY TEST

	Words misspelled	Percent of college freshmen missing word	Percent of 8th grade students missing word*	No. of ways in which freshmen misspelled word	Point or points at which difficulty was encountered	
1.	place	.6	0	1	ac (100)†	
2.	buy	.9	2	1	u (100)	
3.	advanced	.9	3	3	ced (100)	
4.	color	1.1	4	2	o (75)	
5.	attempt	2.6	8	2	tt (83)	pt (17)
6.	affair	2.6	9	3	ir (83)	
7.	superior	2.9	13	3	per (83)	
8.	waste	2.9	12	3	ste (100)	
9.	counting	3.7	2	3	u (33)	ing (67)
10.	amount	4.3	1	1	m (86)	
11.	certain	6.0	6	6	er (21)	ai (68)
12.	convinced	6.0	19	8	inc (86)	
13.	apply	8.0	4	1	pp (100)	
14.	employee	8.0	25	3	ee (100)	
15.	believe	8.6	9	4	ie (86)	
16.	neighborhood	9.1	15	10	eigh (71)	or (17)
17.	capital	9.7	16	4	al (94)	
18.	except	10.3	5	3	ex (91)	
19.	beginning	11.1	18	3	nn (94)	
20.	sincerely	12.3	31	8	cere (100)	
21.	choosing	13.4	19	7	oos (98)	
22.	excitement	14.6	24	6	cite (98)	
23.	considerably	14.9	30	19	s (37)	ably (37)
24.	performed	15.1	6	4	er (98)	
25.	ample	15.4	3	4	le (98)	
26.	administrator	15.4	29	13	is (47)	or (27)
27.	lining	15.7	19	2	ni (100)	
28.	advice	17.4	16	2	c (98)	
29.	circuit	18.9	45	8	ui (95)	
30.	principal	20.9	21	4	al (94)	
31.	prior	22.9	52	15	ior (88)	
32.	appendicitis	22.9	58	14	pp (31)	is (42)
33.	materially	24.6	46	8	ial (96)	
34.	receipt	26.6	34	10	eipt (99)	
35.	permanent	26.9	38	18	ane (91)	
36.	advisable	26.9	25	5	sa (100)	
37.	affectionately	27.1	53	18	ate (88)	
38.	ingredients	29.4	56	14	in (76)	ients (14)
39.	ninety	30.3	23	3	e (99)	
40.	mortgage	31.1	40	5	tg (100)	
41.	assessment	31.7	42	13	ess (85)	
42.	grateful	35.4	49	4	ate (97)	
43.	interfere	36.6	22	6	ere (99)	
44.	losing	36.9	39	4	osi (100)	
45.	lieutenant	37.1	41	23	ieu (78)	
46.	anniversary	38.9	65	12	nn (44)	ary (29)
47.	calendar	39.4	36	7	endar (99)	
48.	privilege	45.1	56	17	ile (91)	
49.	accommodate	60.3	61	7	mmo (98)	
50.	phosphorus	75.1	72	9	rus (96)	

* See the Iowa Spelling Scales, Grade VIII.

† The number in each parenthesis denotes the percent which the number of students who misspelled a word at a given point is of the total number of misspellings of the word.

for each individual were identified. Learning could then be focused on individual needs.

Analysis of spelling difficulties comprised the third step in the remedial project. The analysis was made on the assumption that weaknesses in learning can be most effectively and economically overcome in proportion as learners are able to identify the difficult parts of troublesome words and are then able to focus learning at these points of difficulty. Each student readily made his own analysis by underlining the words in his mastery list at the point or points at which they were misspelled in the diagnostic test. Table III, which is based on the results of the preliminary test, shows how this phase of the remedial program worked out in actual practice.

Several facts of significance are revealed by a study of the original data from which Table III was constructed. First, certain words presented only one point of difficulty for the great majority of freshmen who misspelled them, while in other words a considerable percentage of students encountered difficulty at one of two or more points. Examples of this fact are afforded by the words *lining* and *appendicitis* (see numbers 27 and 32 in Table III). Every student who missed the word *lining* misspelled it by doubling the letter *n;* on the other hand, of the freshmen misspelling the word *appendicitis,* 31 percent failed to double the letter *p* and 42 percent encountered difficulty with the last syllable. Second, several words were missed by a greater percent of college freshmen than by eighth grade pupils. Thus, the word *ample* (see number 25 in Table III) was missed by 15 percent of the freshmen and by only 3 percent of eighth graders. Third, one is amazed at the number of different ways in which college entrants misspelled some of the most commonly used words. The number of different misspellings per word for the fifty words in the preliminary test ranged from 1 to 23. The average number of ways in which the various words were misspelled was 7. One-fourth of the words were misspelled in ten or more different ways. The word *calendar,* which received the average number of misspellings, will illustrate the ingenuity with which college freshmen attempt to spell a word which they have not mastered. The number in each parenthesis denotes the number of freshmen who attempted to spell the word in the way indicated. The misspellings were as follows: calender (59); calander (44); calandar (29); calandor (3); calendor (1); callander (1); callendar (1).

Mastering the words which were missed in the diagnostic test constituted the fourth step in the remedial program. In this phase of the

work each freshman poor in spelling studied his own troublesome words. Care was taken to supply the student with an effective technique of word mastery by encouraging him in his independent study to observe the following suggestions: (a) pronounce the word correctly and look closely at each syllable as you say it distinctly; (b) close your eyes and try to recall how the word looks syllable by syllable as you pronounce it in a whisper; (c) open your eyes to make sure that you were able to recall the correct spelling; (d) repeat the two preceding suggestions as often as you need to in order to fix the correct form of the word; (e) write the word without looking at it just as soon as you are sure that you can recall the correct form, then check; (f) repeat the preceding suggestion two or three times without looking at the correct form, then check; (g) mark the word in your mastery list if it is especially difficult for you.[1] After the student had been equipped with his individual word-list and with a technique for word mastery, he was made to realize that mastery was his individual responsibility. He was also encouraged to set a time when he would be ready for a mastery test.

The final step in the remedial program consisted in giving a mastery test, which was equivalent in difficulty to the preliminary test. The purpose of the retest was to make sure that a high degree of mastery had been attained and to obtain an objective measure of improvement. The passing score for the retest was set three points higher than that for the preliminary test in order to counterbalance the loss in ability which is likely to occur after intensive study has ceased. If a student failed to pass the retest, he was required to go back to work on his mastery list. Table IV shows the amount of improvement which was made.

Of the 201 freshmen who were assigned to remedial work in spelling during the early part of the school year, sixteen withdrew from college. The remaining 185 have been retested. Of this number, 180 have made a satisfactory passing score and have been released from further remedial work in spelling.

An analysis of the data contained in Table IV shows that the great majority of the freshmen who remained in college made marked improvement in ability to spell the 2936 words in most common use. Nineteen students made perfect retest scores. All of the students, except

[1] Horn, Ernest. "Principles of Method in Teaching Spelling As Derived from Scientific Investigation." *Eighteenth Yearbook of the National Society for the Study of Education*, Part II (1919), pp. 52-77.

three, made higher scores on the retest than the best score which was made by any of them on the preliminary test. All of the remedial students, except two, attained the median score for college freshmen; all, except five, attained the median score for teachers in service. The mean point gain of retest over pretest score was 11.9, which represents a percentage gain of 32.96. The individual point gains are even more striking. The smaller gains shown in the table need to be interpreted in the light of the fact that the better spellers were confined within a narrower range of improvement than were the poorer spellers.

TABLE IV

IMPROVEMENT IN SPELLING MADE BY 185 COLLEGE FRESHMEN

Preliminary test scores		Retest scores		Individual points of gain	
No. of words right	Frequency	No. of words right	Frequency	Points of gain	Frequency
42-43	36	50	19	32-33	1
40-41	35	49	45	30-31	2
38-39	29	48	64	28-29	1
36-37	19	47	52	26-27	1
34-35	6	45	1	24-25	6
32-33	17	44	1	22-23	3
30-31	13	43	1	20-21	12
28-29	12	41	1	18-19	12
26-27	6	38	1	16-17	10
24-25	3			14-15	18
22-23	5	Total	185	12-13	8
20-21	1			10-11	22
18-19	1			8- 9	32
16-17	2			6- 7	44
				4- 5	13
Total	185			Total	185
Average	36.1*		48.0		11.9
75 percentile	41.5		49.4		16.4
50 percentile	38.7		48.6		10.3
25 Percentile	32.6		47.8		7.6

* All computations were derived from ungrouped data.

The following statements, which are confirmed by the data which have been presented, are made by way of summary and conclusion.

1. College freshmen vary widely in ability to spell the words in most common use. Some of them spell better than the best one-tenth of teachers in service, while others do not possess seventh grade ability.
2. Marked improvement in spelling may be expected from a technique of learning which first helps each student to discover his own difficult words and then enables him to identify the crucial point or points in each troublesome word.
3. Poor spellers need to be provided with a technique of word mastery.

4. Different students may encounter different points of difficulty in the same word.
5. Study and practice should be focused on the part or parts of words which prove most difficult for the individual students.

THE OBJECTIVITY, RELIABILITY, AND VALIDITY OF AN ESSAY EXAMINATION GRADED BY RATING

VERNER MARTIN SIMS
University of Alabama

INVESTIGATIONS of the objectivity of the essay examination graded in terms of percentage are plentiful.[1] The results of these studies are varied; but Ruch, after tabulating 285 reported correlations between different gradings of the same papers, and finding the median correlation coefficient reported to be .59, summarizes as follows: "In view of all the facts, we must conclude that the central tendency of the old-type [essay] examination is toward a reliability [objectivity of scoring][2] not far from .60 to .65; it is certainly less than .70 on the average."[3]

Such subjectivity would seem to suggest that the grading of an essay examination is scarcely more than a matter of estimating the performance of the testee, consequently the writer decided to apply a commonly used rating technique to their grading and attempt to determine its value. Assuming the possibility of using such a technique its advantages are obvious. Such grading would involve much less time than conscientious percentage grading; and, what is even more important, it would tend to reduce the seeming accuracy of measurement to something comparable with the true accuracy of the procedure.

The general purpose of the problem reported here was to determine the objectivity, reliability, and validity of essay examinations when rated. Practically all of the studies made on essay examinations have been interested in the objectivity of scoring; that is, they have been concerned with the extent to which competent judges agree as to the score which a particular performance should receive. These investigators have, possibly rightfully, assumed that if the examinations are not objective it is useless to investigate their reliability and validity. A good examination, however, must meet the three criteria; and unless one has a preconceived

[1] For good summaries of many such studies see: Ruch, G. M. *The Objective or New-Type Examination* (Chicago: Scott Foresman & Co., 1929), Chapter III.

[2] Ruch uses the term *reliability* to indicate either objectivity of scoring (agreement of competent judges) or adequacy of sampling (agreement of two forms), but I take it that in the reference quoted he refers to objectivity of scoring.

[3] *Ibid.*, p. 106.

notion as to the value of an essay examination, a well-rounded investigation should include not only evidence of objectivity but also evidence of reliability and validity.

PROCEDURE

The general procedure used in this study was to prepare and administer to the same group of students two equivalent forms (Form A and Form B) of an essay examination and an objective test covering the same subject matter. These examinations were based upon Dashiell's *Fundamentals of Objective Psychology,* Chapters IX (Motivation), X (Postural Responses), and XI (Intelligent Behavior), and were administered to a group of 80 students who had just completed the unit of work. No attempt was made to cover the same items in the two types of tests. They represented two independent samplings of the material, having been prepared at different times, and because of the nature of the tests differed radically in the selections made from the total material.

Each form of the essay examination consisted of ten questions. The forms were made by dividing twenty questions at random. Fifty minutes working time was allowed for each form, and all of the students except one finished in the allotted time. The questions were typical essay questions as the following samples indicate:

Form A, No. 1. What is the significance for the problem of learning of (a) motivation, (b) attention (c) intelligence?

Form B, No. 2. What is the relation between motives and habits?

Form A, No. 5. What is attention? Why is it of importance to the psychologist?

Form B, No. 7. Make one statement concerning each of the following: (a) duration of attention (b) effect of distractions on concentrated work.

Form A, No. 10. Of what value are intelligence tests in the schools?

Form B, No. 10. Of what value are intelligence tests in the vocational world?

The objective test consisted of 40 true-false statements and 34 completion elements. Fifty minutes was allowed for completing it, but most of the students finished in slightly less time. The average working time was approximately 35 minutes. The coefficient of reliability based upon 80 cases and determined by correlating split halves and stepping up this coefficient by means of the Spearman-Brown Prophesy Formula was .83.

For convenience in grading, the essay papers were divided at random into four sets of twenty examinations each, and the sets were graded (and are here considered) separately. Five readers graded both Form A and Form B of *Set I*, five graded both forms of *Set II*, six graded both forms of *Set III*, and three graded both forms of *Set IV*. The papers were without names, numbers being used to designate the students, and no reader knew that Forms A and B were papers written by the same student. The readers were: two instructors in general psychology, two student assistants in the Psychology Department, and fifteen students taking the course in general psychology. The instructors, of course, had had experience at grading essay papers. The student assistants had never graded an essay examination in general psychology before. Only students having an A or B record in the course were used as readers, but eight of them had never before graded an examination paper and none of the fifteen had ever graded a psychology paper. The readers were instructed to use the text in working out for themselves satisfactory answers to the questions and then to grade the papers according to the following directions:

1. Quickly read through the papers and on the basis of your opinion of their worth sort them into five groups as follows: (a) very superior papers, (b) superior papers, (c) average papers, (d) inferior papers, (e) very inferior papers. (Remember that in a normal group you would expect to have approximately 10 percent of *very superior* and 10 percent of *very inferior* papers, 20 percent of *superior* and 20 percent of *inferior* papers, and 40 percent of *average* papers. Do not, however, try to conform rigidly to this rule. Your group may not be a normal one.)
2. Reread the papers in each group and shift any that you feel have been misplaced.
3. Make no attempt to give numerical grades or to evaluate each question. Place each paper on the basis of your general impression of the total.
4. Assign letter grades to each group; beginning with A for the *very superior* group, B for the *superior* group, etc., and report the grade for each student. (Report grades on a separate sheet. Make no marks on the examination papers.)

The objective test was given to all of the students in the Department who were taking general psychology (nine sections of approximately 40 students each studying under four different instructors), and on the basis of the distribution of these scores grades were assigned as follows:

Sigma Range			Letter Grade
1.51σ +			= A
$.51\sigma$ to		1.5σ	= B
− $.5\ \sigma$ to		$.5\sigma$	= C
− $.51\sigma$ to	−	1.5σ	= D
− 1.51σ −			= F

Using this key, the scores on the objective test for the 80 students in the study were converted into grades.

By means of calculating coefficients of correlation between the various grades on these measures evidence of objectivity, reliability, and validity was secured. Since the grades for both types of examinations were reported in terms of letters the product-moment formula is cumbersome to use for calculating coefficients of correlation. Holzinger suggests, instead, the use of the coefficient of contingency.[4] All of the coefficients reported here are calculated by means of this formula. The most serious objection to its use is that in a five-fold classification, such as we have used, a perfect relationship is not expressed as +1.00, but as +.894. If this is borne in mind, however, in interpreting the coefficients, it will suffice.

OBJECTIVITY

Objectivity is used as meaning the extent to which competent judges agree as to the grade that a particular performance merits. Competent judges are defined as persons knowing the subject matter and knowing the rules for grading. Since graders of objective examinations are always provided with a key and explicit directions for scoring, this is the same usage as is observed with objective tests.

The objectivity of the essay papers graded by rating was determined by correlating the grades of each reader with the average of the grades given by all other readers of a particular set of papers. Table I presents these coefficients for each form of each set. The table is read thus: the grades given by the first reader of *Set I, Form A*, correlates with the average of the grades given by the other readers of this form .83; the grades given by the first reader of *Set I, Form B*, correlates with the average of the grades given by the other readers .70; the grades given by the first reader of *Set II, Form A*, correlates with the average of the

[4] Holzinger, Karl J. *Statistical Methods for Students in Education* (Boston: Ginn and Co., 1928), pp. 273-278.

grades given by the other readers of this form .83; etc. (It will be noted that the same person read Form A and Form B of each set but the readers for the various sets were not the same persons.)

TABLE I

SHOWING THE CORRELATION BETWEEN GRADES GIVEN BY EACH READER AND AVERAGE OF GRADES GIVEN BY OTHER READERS

	Set I		Set II		Set III		Set IV	
Reader	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
1	.83	.70	.83	.78	.78	.77	.81	.70
2	.81	.79	.78	.75	.77	.76	.76	.82
3	.82	.70	.78	.75	.75	.79	.72	.81
4	.81	.74	.79	.80	.76	.76		
5	.79	.78	.78	.80	.69	.76		
6					.82	.77		

For purposes of interpretation, since most readers are, perhaps, unfamiliar with the coefficient of contingency, the distribution tables for three selected coefficients (the highest, the lowest, and an average) from this group are presented in Chart I.

CHART I. SHOWING DISTRIBUTION TABLES FOR COEFFICIENTS OF CONTINGENCY OF .69, .75, AND .83

	F	D	C	B	A
A					
B			1	3	1
C			6	2	1
D		1	2		
F	2	1	1		

C of C = .69

	F	D	C	B	A
A			1	1	1
B			4	2	
C		2	2		
D		1	3		
F	2	1			

C of C = .75

	F	D	C	B	A
A					1
B			1	2	
C		1	7	1	
D	1	1	1		
F	3				

C of C = .83

Although these coefficients vary a great deal and are not extremely high, they are (accepting Ruch's statement) superior to those found when examinations are graded by percentage. The average coefficient found is .77, and the range from highest to lowest is only 14 points. The probable error of the coefficients was not found, but this range of 14 points is almost surely less than five times the average probable error.

RELIABILITY

Reliability is here used to indicate the consistency of the performance of the testee. Obviously the subjectivity of the grading will reduce the reliability, but with the grading system used what agreement is there between the grades given by a particular reader on equivalent forms of

a test? This reliability was determined by correlating the grades given by each reader on Form A with those given by the same reader on Form B. In addition, the coefficient for the average grade given on Form A with the average grade given on Form B was determined for each set. These last coefficients might be interpreted as the reliability that would tend to be found as the subjectivity of the grading was reduced.

These coefficients are presented in Table II, and, for comparative purposes, the coefficients of reliability for the objective test used are also presented. These coefficients were determined as follows: the objective test was split into random halves and, on the basis of the distribution of the scores on the half, letter grades were determined for each half by the same procedure as that used above for the total test. The letter grades made by the students in each set on the two halves were then correlated by means of the coefficient of contingency and this coefficient stepped up with the Spearman-Brown formula. Holzinger states in a private communication that this stepped-up coefficient may be interpreted as the minimum coefficient that would be found.

TABLE II

SHOWING THE CORRELATION BETWEEN GRADES ON FORM A AND GRADES ON FORM B FOR EACH READER AND FOR THE AVERAGE OF ALL READERS

Reader	Set I (A·B)	Set II (A·B)	Set III (A·B)	Set IV (A·B)
1	.67	.72	.69	.73
2	.73	.74	.73	.71
3	.69	.76	.71	.69
4	.69	.73	.74	
5	.73	.67	.79	
6			.67	
Average Grade All Readers	.70	.73	.73	.71
Objective Test Stepped-up Reliability Coefficient	.78	.83	.74	.87

As would be expected, since lack of objectivity reduces the reliability, these coefficients are lower than these found for objectivity; but they are still relatively high, the range being from .67 to .79 and the average coefficient (.72) being but nine points less than the average reliability coefficient (.81) found for the objective test.

VALIDITY

Evidence of validity is not so easily obtained. For purposes of this study we will define validity as the extent to which the essay examination measures the same thing that an objective test over the same subject matter measures. Perhaps the confirmed addict to essay examinations

would consider this doubtful validation, but it seemed desirable to determine the extent of agreement between the two types of tests. The *standard* nature of the grades on the objective test, due to their being based upon the distribution of scores made by approximately 450 students studying under various instructors, gives added meaning to such a procedure.

The correlation between grades on the objective test and grades given by each reader of each form of the essay test are presented in Table III. The correlation between the average grade for a form and grades on the objective test, and the average of both forms and grades on the objective test, is also presented for each set. These last coefficients should be interpreted as indicating the correlation that would tend to be found as the objectivity and reliability of the essay test was improved.

TABLE III

SHOWING THE CORRELATION BETWEEN GRADES ON THE ESSAY EXAMINATIONS AND GRADES ON THE OBJECTIVE TEST

	Set I		Set II		Set III		Set IV	
Reader	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
1	.72	.61	.77	.69	.75	.70	.68	.73
2	.67	.63	.69	.70	.70	.71	.62	.66
3	.63	.62	.74	.75	.71	.75	.67	.66
4	.65	.66	.69	.78	.71	.68		
5	.66	.68	.73	.81	.81	.68		
6					.69	.65		
Average All Readers	.67	.70	.76	.82	.75	.74	.65	.72
Average All Readers, Forms A and B Combined	.73		.78		.82		.75	

Since neither of the examinations used has perfect reliability one would not expect to find perfect agreement. The maximum correlation expected would be equal to the square root of the products of their reliability coefficients. Using the average reliability coefficient found for the essay test (.72) and the stepped-up coefficient (.81) for the objective test the maximum possible correlation would be $\sqrt{.81 \times .72} = .77$. The coefficients for the individual readers range from .62 to .81 with an average of .70. We thus see that within the reliability of our measures there is high agreement between the two types of examinations. (In passing it should be mentioned that this high agreement between the essay and objective test, coupled with the greater reliability and objectivity of the objective test, might well be interpreted as meaning that the objective test is a better measure of the same thing that the essay examination measures.)

SUMMARY

An essay examination graded by rating was found to have an average coefficient of objectivity of .77 which is considerably higher than the average coefficient reported by Ruch for essay papers graded by percentage.

The average coefficient of reliability found was .72, which is approximately ten points lower than the reliability of an objective test over the same subject matter and demanding slightly less testing time.

The average correlation found between the essay examination and an objective test over the same subject matter was .70, which is within seven points of the correlation that the reliability of the tests would permit.

The evidence presented here is too meager to permit the drawing of general conclusions, but the findings are promising. The objectivity of the scoring is certainly not yet what one would desire. The relatively high agreement (as compared with the objectivity) between the grades on two forms would suggest that the unreliability was principally due to subjectivity of grading rather than to limited sampling. The high correlation between the essay grades and grades on an objective test would seem to indicate that mastery of subject matter rather than such things as grammar, punctuation, spelling, handwriting, etc., is being measured.

The relative nature of the grading means that its use is limited. To take an individual from one group and place him in another would mean that his position and consequently his grade would change. The method could scarcely be used for the comparison of two groups unless the papers were pooled together and graded. But, for ordinary examination purposes, if one is convinced that essay examinations test for abilities which objective tests do not reach, or develop abilities which they do not develop, it would seem that such a grading method is worthy of further consideration. The writer is at present experimenting with the possibility of improving the measurement through refining the rating technique.

SHORTER ARTICLES

A COMPARISON OF THREE METHODS OF COMPUTING COMPOSITE SCORES

EDWARD F. POTTHOFF
University of Illinois

IT is a well known fact that a variety of methods are now available for representing a pupil's attainment on a group of tests by means of a single composite score. The relationships which exist among various types of scores, however, are not always clearly understood. It is the purpose of this article to consider the nature of three types of scores, all of which are based upon the same test and involve the same group of students.

The individuals concerned in this study consist of six hundred students to whom was administered the 1924 edition of the Psychological Examination which is edited by the National Research Council. This Examination consists of nine separate tests, each of which yields its own score. The examination provides that a composite score be determined in the following manner. Percentile ranks are computed for each test so that each of the nine raw scores earned by a student may be converted into its corresponding percentile rank. The median of the nine percentile ranks is then found, this median being used as a single index of attainment in the entire examination.

A second method of computing a composite score is to find the total of the raw scores earned by an individual on the nine tests. Under this procedure the various tests have equal weights in determining the composite score only when the variabilities of the tests are equal or approximately equal. The tests under consideration do not have equal variabilities, however, since their standard deviations for the group studied range all the way from 3.34 to 18.65, this large range being due chiefly to the fact that "perfect" scores on these tests range from 15 to 150. A totaling of the raw scores on these tests does not, therefore, result in giving them all equal weights.

A composite score which does give equal weights to all of the tests may be obtained by converting the raw or original score on each test into the corresponding standard score, and then finding the sum of the standard scores earned by an individual on the nine tests, the resulting total being used as a single index of attainment in the entire examination.

This method, of course, is an extremely laborious one, but it is also commonly regarded as the most accurate.

For purposes of this study each of the three types of scores just described was computed for each of the students tested. These scores, hereafter, will be referred to as the percentile scores, raw scores, and standard scores, respectively. The standard scores were modified so as to avoid negative quantities and large numbers. This modification, of course, does not affect the results in any way.

Comparisons among the three types of scores were made in two ways, first by representing the distributions of the scores in a graphic manner, and, second, by computing coefficients of correlation and correlation ratios. Figure 1 shows the distributions of the scores of the six hundred students when the percentile scores, the standard scores, and a normal distribution[1] are used, thirteen class intervals being employed along the base-line in all cases. The raw scores are not shown because the distribution in this case is very similar to that of the standard scores.

It is evident at once from Figure 1 that the distribution of the standard scores approximates the normal distribution much more closely than does that of the percentile scores. The latter tends to be rectangular, all of the frequencies from the third to the eleventh class intervals inclusive falling within the range 49 to 67. Such a result is to be expected, of course, because of the way in which the percentile ranks are calculated; the rectangular distribution is the result of the mechanics of computation, and, therefore, does not necessarily give a correct representation of the distribution of the traits measured. It has been found that psychological traits distribute themselves normally rather than rectangularly when large unselected groups are used, and in view of this fact it appears that the standard scores are more nearly representative of the facts. At least the distribution in this case is much the more normal in appearance.

In order to determine whether or not there were any other significant differences among the three types of scores studied, a number of coefficients of correlation and correlation ratios were computed. The correlation between the raw scores and the standard scores was .952 ± .003. The coefficient between the percentile scores and the standard scores was .940 ± .003, and the correlation ratios were .948 ± .003 and .951 ± .003. The coefficient between the percentile scores and the raw scores was .908 ± .005, and the correlation ratios were .916 ± .004 and .927 ± .004.

[1] The normal distribution was obtained with the aid of the table on "Areas and Ordinates of the Normal Probability Curve," found in *Holzinger's Statistical Tables for Students in Education and Psychology*, pp. 70-71.

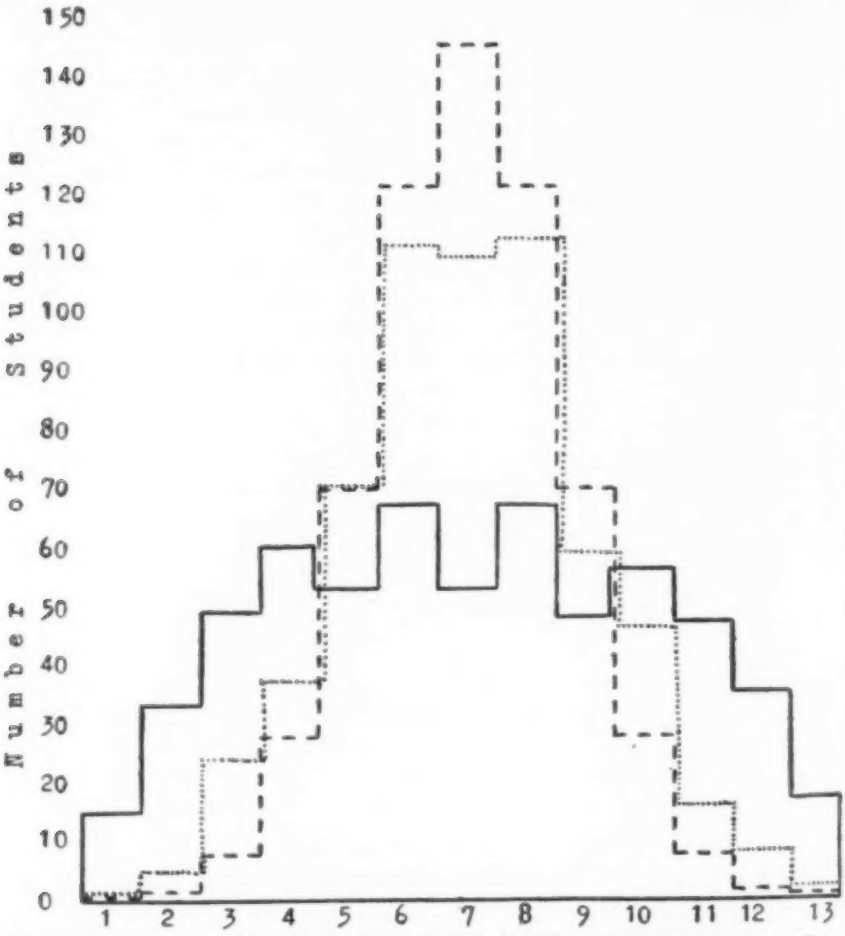


FIGURE 1. SCORES OF SIX HUNDRED STUDENTS DISTRIBUTED ACCORDING TO: PERCENTILE SCORES, STANDARD SCORES, AND A NORMAL DISTRIBUTION. THE CONTINUOUS LINES REPRESENT THE PERCENTILE SCORES, THE DOTTED LINES THE STANDARD SCORES, AND THE BROKEN LINES THE NORMAL DISTRIBUTION

When Blakeman's test for linearity is applied in the case of the relationships found between the percentile scores and the standard scores, the results are 2.22 and 2.63. When the same test is applied in the case of the percentile scores and the raw scores, the results are 2.21 and 3.38. It appears, therefore, that the relationships in both cases tend to be curvilinear rather than linear.

In view of the fact that the correlation between two sets of equivalent scores is frequently found to be as high as .98, it appears that the coefficients of correlation found among the three sets of scores used in this study were relatively low. It should be noted, however, that the coefficients of reliability of our intelligence tests are usually not above .90 and are rarely higher than .95. It is safe to assume then, that the reliability of the Psychological Examination is not higher, and is probably lower, than the coefficients of correlation found among the percentile, raw, and standard scores. In other words, any one of these three methods of scoring is probably more accurate, when either of the other two methods is used as a criterion of accuracy, than is the test itself as a measuring instrument. It follows, therefore, that the three types of scores may be considered to agree very closely with one another and that they may be used interchangeably with chance of only slight error. It should be remembered, however, that because of the nature of the percentile scores, their distributions and their relationships to other sets of measurements may be affected in the ways already shown.

THE PSYCHOLOGY OF THE TABLE OF CONTENTS

J. Howard Stoutemyer
Onargo, Illinois

Professional journals render an invaluable service both to those in the profession and to those in training. This service, however, is not commensurate with their real worth for their perusal is over-limited. Will the proper placement and arrangement of the table of contents increase this perusal? The following data attempt an answer.

I. THE PLACEMENT OF THE TABLE OF CONTENTS

The placement of the table of contents in fifty-five journals in education and closely allied fields is as follows: 35 on the outside front cover, 1 on the outside back cover, 3 on the inside of the front cover, 14 somewhere preceding the reading section, and 2 following the reading section. Of the 20 that place the table inside, 6 place leading articles on the front cover, 4 place pictures or designs, 2 use it for advertising, and the remainder, for the board of editors.

What is the effect of the placement of the table of contents on the outside of the cover on regular or casual readers? To secure an answer, I wrote to twenty educational journals, and received eighteen replies. They may be briefly summarized thus: 1. Only one had definite data. When this journal substituted a few leading articles on the front cover

for the board of editors, favorable comments were received. 2. Present usage is justified on long continued usage. The law of habit would make this effective for regular readers. 3. Each editor expressed his theory. A few types follow: One felt that an illustration or design on the cover was necessary to attract attention to the contents of the journal. "We believe the publication of the table of contents on the cover is a well-established practice." "We *believe*, if we cannot *prove*, that the custom is helpful." "My general impression is that readers prefer the table of contents on the front cover." "I have no data, but, theoretically, I see definite advantage."

Through the courtesy of the *Journal of Educational Research*, my questions were sent to its subscribers. Of the fifty replies, forty-seven were very heartily in favor of the present practice, two felt it was an idle question, and one preferred a few leading articles and the full table inside. A third of the forty-seven stated: "I like it," "I heartily commend the policy." About half stated: "It saves time," or "It is much easier to see whether the articles are of a nature to appeal to me." Three mentioned the greater ease in looking over back files when they sought materials on a problem. Two librarians found it aided in finding references for students. They report that casual readers read more than they would otherwise, for a flashy title or a well-known name attracts their attention. While this number of replies is too limited to warrant general conclusions, we can state that this group favors the placement of the table of contents on the front cover because it saves time, it aids in selecting topics of interest, and it aids in looking over back numbers. General magazines, competing with two hundred others for public patronage at news stands, find an attractive cover necessary for effective news stand appeal. Since professional journals have a relatively small news stand sale, it would seem that cover designs should be for the convenience of subscribers rather than for the casual purchasers.

From a number of experiments on the placement of advertisements in magazines, Starch[1] concludes that the outside covers have about three times the value of inside positions, though the pages next to the covers and those facing the reading section are preferred inside places. From a composite graph of the studies of Adams, Starch, and Strong, the back cover has about two-thirds the drawing power of the front cover and the intermediate pages less than half.[2] Starch explains these data on

[1] Starch, D. *Advertising*, p. 112.
[2] Adams, H. F. *Advertising and its Mental Laws*, p. 222.

the grounds of (1) primacy, the first part of the magazine to be seen, (2) frequency, the greater number of times it is seen, and (3) vividness, for there is less distraction from competing stimuli. Then, too, the right-hand page has nearly twice the drawing value as the left. Since five of these magazines used left-hand pages, their tables of contents would not have full drawing power. As an advertisement of the contents of the journal, placement on the front cover has many advantages. The next best place seems to be the first right-hand page inside the cover.

II. LEGIBILITY OF THE TABLE OF CONTENTS

Ease of comprehension is a chief factor in securing attention, and this depends on the legibility of type. Legibility depends on the form of the letters, size, heaviness of face of type, and width of the white margin surrounding it. "Certain faces of type are much more legible than other faces," and "the relatively heavy-faced types proved more legible than the light-faced types."[3] For example, bold faced Cheltenham can be seen at 286.2 centimeters as against 259.6 for the Italic. Although this showed a loss of one-eighth, there were several journals that use the Italic for titles. According to the judgments of two hundred students, nearly half of these magazines could increase the effectiveness of their table of contents by using more legible type. The same conclusion would hold true of the spacings between words and titles. No doubt color makes the journal more attractive; but great care must be exercised to secure a proper contrast between the color of the cover and the type. One magazine uses a light gray cover and a light Italic type, another uses a dark brown cover and a dull black ink. In neither case is the contrast enough for legibility.

Lines are sometimes used effectively to separate portions of the contents, such as the title of the journal at the top, the table of contents in the body of the cover, and the publisher at the bottom. Lines are also used for borders which serve the purpose of a frame of a picture by shutting out competing stimuli. An effective border also secures unity of impression where many different headings are displayed.

III. THE PLACEMENT OF THE NAMES OF CONTRIBUTORS

A small number of magazines either place the author's name first or else print it in a larger or bolder type. In general, readers look first

[3] Rothlein, Barbara Elizabeth. "Legibility of Different Faces of Printing Type," *American Journal of Psychology*, XXIII (1912), pp. 1-36.

for the topics and then for the author; hence the topics should be made prominent by position and legibility.

Although general practice places the title and then the author's name in the body of the magazine, there are a number of journals that place the author's name following the article. There seems to be a distinct gain in favor of the practice of placing the author's name immediately following the title, for it saves looking back to the table of contents or to the close of the article.

During its latter years, the *Educational Review* placed a short introductory item about the author immediately below his name. For journals with a large circulation and many new contributors this has real value. For several years the *American Journal of Sociology* has placed a brief abstract of the article immediately following the author's name. This enables the reader at a glance to catch the major problem discussed. Summaries or conclusions at the close of the article are also felt to be worth while.

IV. CONCLUSIONS

1. The placement of the table of contents on the outside front cover of professional journals has much in its favor.
2. The table of contents should have a clear but not over large title.
3. The data of issue, year, and volume should also be easily read.
4. The title of articles should be printed in clear, legible type, followed by the author's name and the pages.
5. When colored covers are used, there should be a sharp contrast with the type used.
6. Lines are helpful in focusing attention and in securing unity in the table of contents.
7. The board of editors and other business matters may well be placed on the inside of the front cover, as a number of journals now do.

EDITORIALS

RESEARCH IN HIGHER EDUCATION

CURRENT interest in problems of curriculum reorganization, methods of teaching, and administrative procedure at the college and university level is to a considerable extent the result of a rapid increase in student enrollment since the late war. This influx of a less select student body, with a wider range of individual differences, needs, interests, and capacities, and the resulting expansion of a limited curriculum and modification of inflexible methods of teaching have introduced perplexing problems for solution in higher institutions.

A large amount of educational literature concerned with such problems has been accumulating. The United States Office of Education is publishing a bibliography of some 13,000 titles in the field of higher education. National educational and learned associations and foundations have been active in the promotion of investigations at the college level. In 1930, eighty-two books, monographs, and bulletins appeared in this field. Numerous surveys of higher institutions have been conducted. To name only a few major studies, there are: the Minnesota experiments with class size and science teaching, the Land Grant College Survey, the Modern Foreign Language Study, the Commonwealth Teacher-Training Study, and the National Survey of Teacher Education. National foundations and societies, denominational groups, regional organizations, state governments, teachers associations in subject-matter fields, and individual colleges and universities have promoted intensive surveys of the higher institutions involved. The educational periodicals dealing with college and university problems now include: *Journal of Higher Education, Junior College Journal, School and Society, Educational Administration and Supervision, Liberal Arts College Bulletin,* and *Bulletin of the Association of American Colleges.*

The foregoing facts imply that a wealth of material is available for use in university courses dealing with problems of higher education. During the summer of 1931 twenty-four universities offered courses in this field for administrators, teachers, and graduate students. There is need for study and understanding of problems in higher education on the part of workers at both college and high-school levels. Secondary-school

principals and teachers who know the entire educational system are able to provide more adequate opportunity and guidance for high-school pupils who expect to enter college. Such courses are of especial interest to candidates for the Doctor's degree in view of the fact that 75 percent of this group after graduation go into college or university work. Many of these students work out theses concerned with college problems. From 1918-1929 there were completed 244 Doctors' theses dealing with problems of higher education. In 1930-1931 there were under way 181 Doctors' theses in higher education.

Does this rather formidable statement of investigational activity at the college level mean that most of the problems in this field have been solved? Not at all. Especially is evidence lacking concerning the merits of certain teaching and curriculum innovations at both graduate and undergraduate levels. The writer has been impressed by the absence of such objective evaluation in collecting from higher institutions material for the Department of Research News and Communications of this journal. At the graduate level, in objective terms, what are the merits of the Iowa plan of individual or independent study, of the Yale seminar organization, and of the new degree programs in education as described in the 1931 yearbook of the National Society of College Teachers of Education? At the undergraduate level, on the basis of experimental evaluation, what is the evidence concerning Rollins' workshop or laboratory plan, Cincinnati and Antioch's coöperative or part-time programs, Harvard and Princeton's general reading and tutorial plans, Stanford's independent study scheme, Swarthmore's honors courses, Wisconsin's Experimental College, and Chicago's divisional reorganization?

Granted that administrative difficulties may seem almost insurmountable and variable factors apparently uncontrollable, the challenge to evaluate these plans on an experimental or factual basis remains for the student of objective measurement. Minnesota's work with class size and science teaching is a good beginning in this experimental program. It must be kept in mind that scientific evaluation does not consist of opinions collected from parents, students, or teachers with respect to the merits of a given plan; neither is it journalistic description of the methods in use; it is not even analysis and criticism based on observation and painstaking rationalization. Of just what scientific evaluation does consist is a long story that has been reasonably well told in current volumes which deal with techniques of measurement, experimentation, statistics, and educational research.

CARTER V. GOOD.

REVIEWS

Address all communications relative to book reviews to Editor E. J. Ashbaugh, Miami University, Oxford, Ohio.

EDGERTON, HAROLD A. *Academic Prognosis in the University.* Baltimore: Warwick and York, 1930. vii + 83 pp.

Like the elementary school and the high school, the college is no longer a highly selective institution. Due to the heterogeneous groups of students who enter higher institutions to-day, a number of problems have arisen: revision of objectives, reorganization of the curriculum, modification of classroom methods, provision for individual differences, and invention of valid techniques of measurement. One of the most vital of these problems is concerned with prediction of success.

Professor Edgerton has conducted a study at Ohio State University to investigate the possibility of a system of "cumulative" prognosis of academic achievement. He used the records of two classes of freshmen who entered in the fall of 1924 and 1925. Each student was given the Ohio State University Psychological Examination at the time of entrance. The complete academic record for each student was secured from the time of entrance until June, 1927. Thus, progressively at the end of each quarter the cumulative marks along with the intelligence-test records of the students were used as means of improving the prediction of future academic performance. The results show that the entrance-test scores correlate .517 with scholarship in the junior college. When the marks of the first quarter are added to the test scores as a prognostic measure, the correlations with junior-college scholarship increase to .882. At the end of the second quarter, the cumulative marks and test scores correlate .932 with academic standing at the end of the junior-college period.

The results of this study are in general agreement with the conclusions of other investigations of prediction of success. Coefficients of correlation between either high-school marks, college-entrance examinations over high-school subjects, or intelligence test scores and college marks may be expected to range from about .40 to .50 or perhaps higher; combinations of these factors will often yield correlations of .60 or above; and, if measures of study and other relevant habits are included, the correlations may be expected to rise as high as .75 or .80.

However, such measures of prognosis are concerned only with prediction of success in the academic work of the university. The investigator who shows objectively and accurately the relation between current instruments of prognosis, including marks in college, and success in life or adequate adjustment to extra-school environment will perform a valuable service. If correlations should prove lower than expected, either the curriculum requires revision to conform more nearly to social need or new instruments of prognosis should be invented to predict other than academic ability, or both.

CARTER V. GOOD
University of Cincinnati.

THORNDIKE, EDWARD L. *Human Learning*, New York: The Century Co.

This is Professor Thorndike's latest book and should prove especially interesting to laymen in that it is a non-technical discussion of certain facts and principles pertaining to the learning process which the author has brought together during many years of investigation in the fields of both human and animal learning.

The first ten chapters are revised lectures dealing with the nature of human learning. With one fell swoop, the author disposes of "adjustments," "configurations," "drives," "integrations," "purposes," "tensions" and the like by saying that "all of these facts seem to me to be reducible so far as concerns their powers to influence the course of thought or feeling or action, to *connections* and to *readinesses* in the nervous system. Learning is connecting. The mind is man's connective system." (Page 122).

The physiological basis for this theory is equally simple and, as the author himself admits, highly speculative, (as indeed are all physiological theories of learning). Each neurone is supposed to live much as would an amoeba, or any one-celled animal, with all movement or changes of position being restricted to its ends. Since a neurone's life processes consist of receiving and transmitting stimuli, any synaptic bond, the conduction across which produced "annoying states of affairs," would be given up, and those which produced "satisfaction" would be maintained. And so we have a restatement of his old hedonistic and association theory of learning, supported, however, by much new experimental data that gives a certain evaluation or weight to his theory. This theory he now restates in the light of this experimental data and the numerous criticisms that have been directed towards it in recent years by Gestalt psychologists and other workers in the field of learning.

The last two chapters of the book are concerned with the evolution of learning and its future possibilities.

"Man does not, so far as we know, have new varieties of neurones or new modes of action of neurones, but he does have many more of them with far greater possibilities of interconnection than exist in other animals." (Page 168). This fact is made the basis for his quantitative theory of learning. This "quantitative difference in associative learning is . . . the producer of the qualitative differences which we call powers of ideation, analysis, abstract and general notions, inference, and reasoning."

The outlook presented by the book is optimistic. For as the author puts it: "We still have much to learn about eugenics," "but even now we know . . . that with selective breeding within 100 generations, we could hope to have a race of men whose average would be much nearer to the intellect of Newton, Pasteur, Gladstone, or Edison than to that of the average man of to-day."

Professor Thorndike's discussion of the nature and evolution of man's learning process explains many fundamental facts and principles of human learning in the light of significant experimental data obtained from investigations carried on during the past three years. These principles and facts are explained in clear and simple language that should be understood by all. The book gives the best statement of the author's theory of learning that he has ever made. It does not, however, present a complete account of the learning process. The establishment of new habits or connections by the action of conditioning stimuli is

only briefly sketched, and the relation of this fundamental type of improvement to the learning process as a whole is not pointed out.

W. F. BOOK

Indiana University.

BETZNER, JEAN. *Content and Form of Original Compositions Dictated by Children from Five to Eight Years of Age.* New York: Columbia University, 1930. v + 53 pp.

This is a study of the original compositions of 1,215 children, ranging in ages from five to eight years. These children, selected from sixty-six classrooms, located in eleven communities in five states, were a cosmopolitan group, Italians, Finns, Russians, and negroes being liberally mingled with Americans. Twenty-nine of the classrooms represented communities of more than ten thousand inhabitants, but thirty-seven were located in towns with smaller populations.

The children dictated their stories to certain competent and sympathetic adults, who were careful to read the compositions back to their young authors as often as desired. After the grossest English errors had been corrected and the compositions themselves had been typewritten, these stories were analyzed to discover similarities and differences in content and form.

Dr. Betzner reaches several interesting conclusions in this study. Children in the first years of school show greater variety than duplication in the choice of themes. They have a greater liking for the realistic story than for any other type, although they do not wholly neglect imaginative, descriptive, rhythmical, and humorous methods of composition. For the most part, children tell stories which deal with first hand experiences, imaginative and literary experiences not being wholly ignored; and the relationships with which they deal are usually normal and wholesome. Compositions from children of five, six, seven, and eight years of age range from nine to 1,074 words, the average story having 66.6 words; they contain from one to sixty-nine thought units, with a median of 7.25; they are usually given titles; they observe time sequence; they have conclusive endings, some repetition to heighten impressions, some characterization, conversation, and unusual phrases; and they pay more attention to motor than to sense aspects.

The study further points out some interesting facts about the increase or decrease of certain characteristics with the progress in age, as well as a comparison of masculine and feminine ability in the young authors.

In the light of this analysis, and after a careful examination of existing textbooks, courses of study, and books on methods, Dr. Betzner comes to the conclusion that education is putting too much stress upon securing uniform products. She believes that the children's willingness to dictate stories to others is an indication that artificial and elaborate devices for securing original compositions are unnecessary. She further believes that too many subjects suggested for composition work accept as a fact the idea contradicted by this study, that a child is most interested in stories in which he himself is the hero. The study indicates also that children are far more interested in narrative than in essay writing, and the author recommends that less of the latter type be required in school.

J. HELEN CAMPBELL

Miami University

MEIER, LOIS. *Natural Science Education in the German Elementary Schools.* New York: Bureau of Publications,

Teachers College, Columbia University, 1930. vi + 158 pp.

Miss Meier's study of the German schools was made during thirteen months spent in actual contact with the schools in which time she also had opportunity to take courses in the changes which are going on in the German system.

The first hundred pages of this volume outline the science work in the elementary schools and in the secondary schools. The last third of the book is devoted to teacher training. The last chapter deals with a comparison of the German and American work in science.

Comparing the science work in the first eight grades of typical states in this country and in the corresponding school level of Prussia, Miss Meier finds that California devotes from 1 to 2 percent of the total teaching time of grades I to VIII to science; Massachusetts devotes from 4 to 8 percent; and New Jersey 3 to 5 percent while Prussia gives from 17 to 39 percent. Thus, California devotes 185 minutes per week to science instruction as the total for the eight grades; Massachusetts devotes 600 minutes, New Jersey 440 minutes and Prussia 1,755 minutes.

There is a similar contrast in the time devoted to science in the secondary schools. The student must take some science every year during the nine years of his secondary school course. This totals 35 year hours of instruction in the science course and 18 year hours in the classical course. On the whole, the girls have more science than do the boys, devoting more time to biology and less to physics.

The subject matter of science instruction in the elementary schools is not very different from that in our own nature study though it does deal more with materials that are of industrial value. The method of instruction in the elementary school is largely through walks, the pupils recording their studies in permanent notebooks. The method of instruction in the secondary schools is quite similar to that of American institutions excepting that the work is presented largely by the demonstration method instead of by the individual laboratory method.

Even for the elementary schools, the German teacher must have taken at least two years of training beyond the German secondary schools, the certificate from which is about equivalent to our junior college certificate. Since science is required throughout the secondary school, every teacher will have had some training in science. Miss Meier contrasts this situation with figures quoted from Miss Bilig's study on the preparation of students enrolled in various teacher training institutions in the United States. Only 54 percent of these students had had any biology, 37 percent having had it in high school; 43 percent had had physics, 41 percent in high school; 42 percent had had chemistry, 41 percent in high school. The percentage in other sciences was considerably lower than for these three.

After completing the course at the teacher training institution, the young teacher undergoes a two year system of training in service under the state department before he receives a permanent appointment as an elementary science teacher.

E. R. DOWNING
University of Chicago.

JAMESON, EMILY D. *Physical Education for the Preparation of General Elementary School Teachers.* New York: Teachers College, Columbia University, 1930. viii + 118 pp.

The elementary school teacher is being expected to teach physical education

along with her other duties. Out of four decades of rapid development of the health and physical education program, this policy is emerging as a desirable practice. Considered by many as a special subject, health and physical education, nevertheless, has few representatives as special teachers in the elementary field. Professional opinion, at present, believes that the classroom teacher should have the responsibility for this as for other subjects.

This being the case, the contribution made by Dr. Jameson reveals an inadequacy of the training programs of the fifty-six normal schools and colleges whose curricula were analyzed. It is shown, for example, that there is little differentiation within these schools between service courses in physical education and professionalized subject matter courses (p. 44); that specific training in appropriate elementary school activities is not generally given (p. 32 and 97); and that there appears to have been no standard practices relative to the classroom phases of the physical education training courses.

Dr. Jameson's study, therefore, provides a challenge to a further and more complete curriculum analysis so very much needed in this field. The value of a study of what certain colleges are now doing is limited by the value of the sources of information. College training curricula do not always reflect accurately the need of the field. The curricula of these particular colleges are shown to be rather haphazardly developed in relation to each other or to the field. For instance, the principle of adapting activities to appropriate age groups is one of the topics stressed in the theoretical parts of the physical education courses. But in the activity courses one is startled to find football, basketball, and tennis taught in profusion! No wonder we have football midget teams and children of tender years playing basketball before their hands are large enough to hold the ball. Dr. Jameson makes a heroic attempt to show, in her concluding chapter, the inadequacies in the training programs but without a curriculum analysis, using school situations and scientific evidence on age capacities, there can be little help given the problem.

Had not Dr. Jameson been writing a dissertation she undoubtedly would have omitted a great amount of inconsequential statistical material. The author has been thorough, perhaps meticulous, in her task, but the material presented is such a small part of the whole program that further evidence is necessary in order to see clearly the situation. One feels, perhaps, that the title of the book is a misnomer. The large problem remains untouched. The author should be encouraged to follow this volume with an entire book written on the thought in the last sentence of the second chapter: "It is hoped that all institutions will come eventually to differentiate at least one year of their physical education so that each course will be especially adapted to prospective teachers of some specific age level."

D. OBERTEUFFER
State Department of Education.

LOWTH, FRANK J. *The Country Teacher at Work.* New York: The Macmillan Company, 1930. xii + 541 pp.

Perhaps the most outstanding characteristic of this book is the evidence, on almost every page, of the author's intimate acquaintance with rural needs. Another compelling feature is his wholesome philosophy of education and common sense suggestions for applying it. One finds himself again and again men-

tally nodding assent to the clean cut and vigorous statements of purpose and method throughout the book. Moreover, the author indulges but little in high sounding phrases. His illustrations are usually drawn from experience; his suggestions are practical, and he meets rural education at its worst—the one-room rural school.

There are some features of the book, however, which seem to merit less commendation. For example, the chapter of some fifty pages on "Keeping Records and Making Reports" has some very excellent suggestions, but a large part of it is taken up by printed forms and standard items. Rural teachers as a rule have very little to do with selecting the form of the records used and hence would have only a very general interest in them except in the particular records they use in their own schools. Likewise, the chapter on "Thrift For Teacher and Pupil" seemingly would be more adapted to an upper grade pupil's textbook than to occupy space in this book which for the most part presents materials both "useful and usable" in the midst of classroom activities. It may be that sheer prejudice prompts the criticism of the general organization of the book, even though knowing the theme "Objectives" runs through the book, one's sense of order is jarred by having a most stimulating discussion vitally related to classroom procedure on "Personal Influence in Discipline" followed immediately by a chapter on how to save money, with its budgets, compound interest tables, etc. This does not say that the thrift idea is not good, but it is not usually yoked so closely in teachers' books with the discussion of essentially schoolroom activities. Other parts of the book show the same tendencies. In one section "Reading" is listed under "Instructional Objectives," the following section is entitled "Realizing Objectives" and covers "Tools and Skills." Certainly "Reading" is a "Tool" subject and is used fully as much to "realize objectives" as "Learning to use the Library." These are of minor importance, however, when compared with the sterling worth of other parts of the book. One might search through many volumes to find more "vitalized" suggestions on "practical citizenship" and on how to teach it than are given in Chapter XII. The chapter on "The Use of the Dictionary" is helpfully suggestive. The same thing is true of the discussion on "Learning to Use the Library," while the chapter listing the various types of educative seat-work, with plenty of explanation for each, certainly should be a mine of ideas to the rural teacher.

On the whole, this book brings practical suggestions in abundance to the aid of the teacher most in need of them. It expresses a wholesome philosophy of education fitted to her situation. It uplifts and points to better things. It inspires, it encourages, it tempts the teacher to try for greater perfection in teaching, and makes for increasing dissatisfaction with the "outgrown shells." Yes, it deserves a place on the bookshelf of the rural teacher.

E. L. RITTER
Iowa State Teachers College.

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to Editor A. S. Barr, University of Wisconsin, Madison, Wisconsin.

BULLETINS

CHRISTOFFERSON, H. C., CLIFTON, J. L., AND WOOD, E. R. "The Learning and Teaching of Geometry," *Report of the Every Pupil Test in Plane Geometry*, December 2, 1930. Conducted by the State Department of Education, Columbus, Ohio.

Contains a statement of the objectives of the Every Pupil Test program, a general summary of results, and an analysis of errors with suggestions for better teaching.

CORNELL, ETHEL L. "The Effect of Trait Differences in Ability Grouping," *University of the State of New York Bulletin*, No. 981, (October 1, 1931), 27 pp.

Discusses trait differences versus individual differences and the relation of trait differences to homogeneous grouping. Reports the results of an experiment testing the feasibility of ability grouping. The author shows that there is a considerable amount of overlapping for the various groups and subjects, and concludes that the variability within each group is such that grouping will not produce sufficient homogeneity to relieve the teacher of the necessity of providing work for the individual at the stage at which he is.

COXE, WARREN W., BAER, JOSEPH A., VAN COTT, HARRISON H., AND WILEY, GEORGE M. "Courses of Study and Curriculum Offerings in Junior High Schools in New York State." Albany: State Education Department, University of the State of New York, 1931.

Contains an analysis of junior high-school courses for the several school subjects and a discussion of progressive practices in junior high-school courses of study.

FOSTER, I. OWEN, CHAIRMAN, AND OTHERS. "A Study in Teacher Supply and Demand in Indiana," *Bulletin of the School of Education, Indiana University*, VII (September, 1930), 77 pp.

Contains materials on the Indiana teacher supply for 1929-30, teacher-training enrollments in Indiana; new teachers for 1929-30, types of schools in which employed, and subject combinations for high school teachers; high school subject combinations as shown in Indiana state school directories for 1927-28, 1928-29, and 1929-30; and summary and conclusions.

HENDRIX, S. GERTRUDE. "Teaching Devices on the High School Level," *University of Illinois Bulletin*, XXVIII (June 16, 1931), 42 pp.

Contains a classified list of teaching devices for mathematics, physical sciences, natural science, social sciences, English, and foreign languages. A brief, descriptive paragraph is given for each device.

HUTSON, PERCIVAL W. "Measuring the Need for Guidance in Western Pennsylvania—a Partial Survey Directed and Prepared for Publication," *University of Pittsburgh School of Education Journal*, VI (May, 1931), 115-154.

Contains three studies relating to various aspects of this problem. The author finds that guidance is needed; that the choices of both boys and girls of the upper and lower secondary schools are seriously in line with social needs as indicated by the present distribution of workers among occupations.

KEELER, L. W. "An Investigation of the Effect of Subject Deficiencies Upon Accomplishment of Students Entering the College of Engineering of the University of Michigan During the Academic Years, 1927-28, 1928-29, and 1929-30," *Bureau of Educational Reference and Research Bulletin*, No. 138 (March 30, 1931), 68 pp.

The author finds that each succeeding year shows an increased percentage of deficiency, that the collective deficiency in trigonometry, geometry, and algebra is greater than that in all other subjects, and that the growth in deficiency is most pronounced and consistent in the case of geometry, algebra, and physics.

LANGVICK, MINA M. "Current Practices in the Construction of State Courses of Study," *United States Department of the Interior Bulletin*, No. 4 (1931), 84 pp.

Contains materials relating to legal provisions which affect curriculum construction, policies of state educational officials with respect to the publication and revision of state courses of study, findings of recent state educational surveys, in relation to the development of the curriculum, plans and techniques employed in the process of curriculum construction, educational values desired for children of elementary school age, and organization of the curriculum in relation to working schedules and plans for promotion.

MARTENS, ELISE H. "Biennial Survey of Education in the United States, 1928-1930," *United States Department of the Interior Bulletin*, No. 20 (1931), 38 pp.

Contains an analysis of the situation relating to exceptional children in the United States, state support and authorization of special classes, the supervision of special classes, city provisions for exceptional children, county provisions for exceptional children, preparation of teachers, mental hygiene, etc.

ROGERS, JAMES FREDERICK. "The Speech Defective School Child," *United States Department of the Interior Bulletin*, No. 7 (1931), 31 pp.

Discusses the kinds of defects of speech found among school children and special work for speech defects, with a summary and bibliography.

SMITH, HENRY LESTER, AND KRUEGER, LEVI MCKINLEY. "The Psychology of Human Relationships—Individual and Social: A Source Book," *Bulletin of the School of Education, Indiana University*, VII (March, 1931), 107 pp.

The readings are organized around the following topics: (1) psychology of the individual: the individual as an end, human nature, and instincts and emotions; (2) psychology of social relations: group behavior, and some problems in social relations; and (3) psychology of national and international relations.

SMITH, HENRY LESTER, AND LITTELL, HAROLD. "The Philosophy of Human Relations—Individual and Collective: A Source Book," *Bulletin of the School of Education, Indiana University*, VII (January, 1931), 326 pp.

Part I contains readings illustrating the philosophy of human relations in early historical times. Part 2 illustrates the philosophy of human relationships as applied to political theory: nationalism, internationalism, cosmopolitanism; materialism; moral philosophy; morality; religion; and education.

SMITH, MAPHEUS. *A Study of the Unsupervised Behavior of a Group of Institutional Children.* University of Kansas, 1931, 40 pp.

A study of the unsupervised behavior of a group of 46 institutional children. Marked individual differences were displayed in the behavior of children. Boys surpassed girls in attacking and in being attacked. Girls surpassed boys in giving. The older group surpassed the younger group in pattern play and in active social behavior.

THOMAS, EARL, REAVIS, W. C., AND OTHERS. "A Study of the Influence of Departmental Specialization on the Professional Attitudes of High-School Teachers Toward Certain Administrative Problems," *Bulletin of the Department of Secondary-School Principals of the National Education Association*, No. 37 (May, 1931), 70 pp.

Contains material on the organization of the curriculum, the size of the class in relation to instruction, pupil failures, the function of the teacher in the class period, teacher rating, faculty meetings, supervision, and the distribution of the principal's time.

EDUCATIONAL ADMINISTRATION AND SUPERVISION

ANDERSON, EARL W., AND ELIASSEN, R. H. "Subjects Taught in High School," *Educational Research Bulletin*, Ohio State University, X (May 27, 1931), 283-289.

An analysis of 25 studies relating to subjects taught by 98,464 high-school teachers, giving the percentage of teachers teaching in the various subjects.

ANDERSON, H. J. "Correlation Between Academic Achievement and Teaching Success," *Elementary School Journal*, XXXII (September, 1931), 22-29.

The results of this study show that the correlations between academic achievement and teaching success are low.

BUCKNER, MABEL A. "A Study of Pupil Elimination in the New Haven High School," *School Review*, XXXIX (September, 1931), 532-541.

An analysis of the causes for pupil eliminations, with recommendations for the improvement of pupil adjustment in high schools.

COXE, WARREN W. "Study of the Elementary School Teacher in New York State," *University of the State of New York Bulletin*, No. 975 (July 1, 1931), 19 pp.

Contains data relating to the general training, professional training, teaching experience, and nationality of elementary school teachers in the state of New York.

FERGUSON, GEORGE O., JR. "Some Factors in Predicting Academic Success in the College of Arts and Sciences, University of Virginia," *University of Virginia Record Extension Series*, XV (April, 1931), 3-13.

An analysis of the records of students at the University of Virginia from which the author concludes that, by excluding about 200 students of those now admitted, the total number passing on only six hours or less might be reduced by one-half.

GARRETSON, OLIVER KELLEAM. "In-Service Training of Teachers in High Schools in Oklahoma," *School Review*, XXXIX (June, 1931), 449-460.

A questionnaire study of the supervisory practices of high school principals in the state of Oklahoma.

HAWKS, LENA JAMES. "Certain Relationships Between Scholarship in High School and in College," *Johns Hopkins University Studies in Education*, No. 15 (1931), 58 pp.

From an analysis of the records of the entering students of three universities and the relationship between marks in high school subjects and scholarship averages in college, the author concludes that the relationship between marks in individual subjects in high school and first semester and first year scholarship in college is .60; between the general high school average and scholarship in college is .70.

OVERN, A. V. "The Length of Service of Minnesota Teachers," *Elementary School Journal*, XXXII (September, 1931), 34-43.

One thousand thirty-six teachers who began teaching in Minnesota in the school year 1919-20 were traced through their teaching service for a period of nine years. The percentages remaining in the teaching profession for the successive years were as follows: 100, 52, 35, 26, 18, 14, 11, 6, 4.

REINHARDT, EMMA. "General Introductory Courses in Education in Teachers' Colleges," *Elementary School Journal*, XXXII (September, 1931), 57-66.

A questionnaire study of introductory courses in Education as reported by 63 teachers' colleges. Such a course is required in 41 of the colleges.

WELTZIN, J. FREDERICK. "Authority of Normal-School Boards to Establish Teachers' Colleges," *Elementary School Journal*, XXXII (September, 1931), 44-52.

A legal study of the authority of normal-school boards. In general, it may be said that, although administrative boards have wide power and broad discretion in the management and control of the institutions under their care, the power to make drastic changes, such as the creation of new schools, etc., must be expressly delegated, either by the constitution or by statute, or must be extremely clearly implied.

WRIGHT, FRANK L. "A Comparison of College and University Graduates with Graduates of Teachers Colleges in Their Preparation as Prospective Secondary Teachers," *Educational Administration and Supervision*, XVII (September, 1931), 453-461.

The author presents, from an analysis of the transcripts of high school and college students graduating from different types of institutions, a summary of their training programs.

EXPERIMENTAL STUDIES OF LEARNING AND TEACHING

CHAPPELL, MATTHEW N. "Inhibition, Facilitation, Learning: Summation of Stimuli," *Psychological Review*, XXXVIII (July, 1931), 317-331.

After giving a critical analysis of the terms inhibition, facilitation, and learning, the author proposes that these various phenomena are adequately explained by the summation of stimuli.

HOLMES, GRACE. "The Relative Legibility of Black Print and White Print," *Journal of Applied Psychology*, XV (June, 1931), 248-251.

An experimental study of the relative legibility of black and white type. The legibility of black type on a white background is 14.7 percent greater than white type on a black background.

HURLOCK, ELIZABETH B., AND NEWMARK, ELLA D. "The Memory Span of Preschool Children," *Journal of Genetic Psychology*, XXXIX (June, 1931), 157-173.

After reviewing previous investigations the author reports the results of the memory span of 20 preschool children in a public kindergarten in New York City. The author finds that a Digit Backwards Test is too difficult for preschool children, but that the memory span for a Digits Forward Test is a good test of general intelligence, the correlation for abstract words being +.55, and for concrete words, +.52.

JAMES, H. E. O. "The Problem of Interference," *British Journal of Psychology:* General Section, XXII (July, 1931), 31-42.

The notion of interference is used frequently to explain certain phenomena of learning. The author examines the processes of interference, and links it with Pavlov's work on inhibition.

PATERSON, DONALD G., AND TINKER, MILES A. "Studies of Typographical Factors Influencing Speed of Reading: VI. Black Type Versus White Type," *Journal of Applied Psychology*, XV (June, 1931), 241-247.

An experimental study of the relative merits of black and white type. The results favor the black type.

VICKERY, KATHERINE. "The Effects of Change of Work on the Work Decrement," *Journal of Experimental Psychology*, XIV (June, 1931), 218-241.

This study attempts to determine whether under certain conditions there exists a work decrement in two types of mental activity as shown by a cancellation test and an equation test. No work decrement developed in the cancellation test, but a decided decrement developed in the equation test.

WARDEN, C. J., AND COHEN, A. "A Study of Certain Incentives Applied Under Schoolroom Conditions," *Journal of Genetic Psychology*, XXXIX (September, 1931), 320-327.

An experimental study of several types of incentives used under schoolroom conditions. The author finds that these incentives, while not especially effective as far as speed was concerned, were effective in improving the accuracy scores of both boys and girls.

PHYSICAL ADJUSTMENT, GUIDANCE, AND WELFARE

HEDGE, JOHN W., AND HUTSON, PERCIVAL W. "A Technique for Evaluating Guidance Activities," *School Review*, XXXIX (September, 1931), 508-519.

The authors describe a guidance procedure and report the results of a comparison of the aspirations of pupils in September and May with their abilities, to determine whether the later compari-

sons represented closer relations than the former.

CURRICULUM RESEARCH

DAVIS, IRA C. "Analysis of the Subject Matter in the Eight Most Widely Used Textbooks in General Science," *School Science and Mathematics*. XXXI (June, 1931), 707-714.

The author presents a topical analysis of the subject matter of eight widely used textbooks in general science. There is a high agreement in the textbooks for the units on air, water, heat, weather and climate, light, magnetism and electricity, energy and machines, plants, the human body, and foods; fair agreement in the units on sound, the solar system, and rocks and soils; and very little agreement in the units on animals and clothing.

KYTE, GEORGE C. "Specific Requirements in the Two-Year Curricula for Training Elementary School Teachers in Representative Teachers Colleges and Normal Schools," *Educational Administration and Supervision*, XVII (September, 1931), 401-416.

An analysis of the requirements for two-year curricula for the training of elementary teachers indicating the frequency with which various courses are required in representative teachers colleges and normal schools.

WILSON, HOWARD E., AND ERB, BESSIE P. "A Survey of Social-Studies Courses in 301 Junior High Schools," *School Review*, XXXIX (September, 1931), 497-507.

The authors present statistical data relating to the frequency with which courses in early American history, later American history, European background, ancient history, world history, geography, civics, Rugg's fusion course, and local fusion courses, are offered in grades 7, 8, and 9 of the junior high school. He finds little agreement in course requirements.

MEASUREMENTS, STATISTICS, AND METHODS OF EDUCATIONAL RESEARCH

DOWNEY, JUNE E. "Inverse Trends in the Pressey X-O Test and the Downey Will-Temperament Test," *Journal of Applied Psychology*, XV (June, 1931), 273-278.

The author comments upon a recent article comparing the Downey Group Will-Temperament Test, the Pressey Cross-Out Tests, and the Brotemarkle Comparison Tests, pointing out that low coefficients of correlation do not necessarily invalidate her own or other tests of this group since they measure different traits.

MEAD, CYRUS D. "An Industrial Arts Lesson Scale," *Educational Administration and Supervision*, XVII (September 1931), 429-445.

A quality scale for judging the merits of industrial arts lessons in the upper elementary and junior high school grades.

ODELL, C. W. "Still More about Scoring Rearrangement or Continuity Tests," *School Review*, XXXIX (September, 1931), 542-546.

A further discussion of methods of scoring continuity tests.

PEARSON, EGON S. "The Test of Significance for the Correlation Coefficient," *Journal of the American Statistical Association*, XXVI (June, 1931), 128-134.

A discussion of statistical procedures for testing the significance of the coefficient of correlation.

RIETZ, H. L. "Comments on Applications of Recently Developed Theory of Small Samples," *Journal of the American Statistical Association*, XXVI (June, 1931), 150-158.

A discussion of a theory of small samples.

THURSTONE, L. L. "Multiple Factor Analysis," *Psychological Review*, XXXVIII (September, 1931), 407-427.

The author presents a statistical procedure for supplementing Spearman's two-factor method of analysis to include a number of general or group factors.

THURSTONE, L. L. "Rank Order as a Psychophysical Method," *Journal of Experimental Psychology*, XIV (June, 1931), 187-201.

The development of a statistical procedure for a shorter method of applying the rank order method.

MISCELLANEOUS

PETERS, CHARLES C., AND STRUCK, F. THEODORE. "Abstracts of Studies in Education," *Penn State Studies in Education*, No. 2 (1931), 96 pp.

An abstract of some thirty-eight masters' and doctors' dissertations prepared by students of Pennsylvania State College.

"Proceedings of the Seventh Annual Educational Conference," *Bulletin of the Bureau of School Service*, College of Education, University of Kentucky, III (December, 1930), 100 pp.

Contains addresses by Hamilton Holt, Carleton Washburne, Thomas H. Briggs, H. L. Donovan, and others.

NEW TESTS

MEIER, NORMAN CHARLES, AND SEASHORE, CARL EMIL. *Meier-Seashore Art Judgment Test.* Iowa City, Iowa: Bureau of Educational Research and Service, 1931.

The test is designed for use primarily in junior and senior high schools, and in art schools.

RESEARCH NEWS and COMMUNICATIONS

Address all research news and communications to Editor Carter V. Good, Teachers College, University of Cincinnati, Cincinnati, Ohio.

Request for Titles of Ph.D. Theses Under Way.—The Journal of Educational Research expects to publish each year in the January number a list of the Doctors' dissertations under way in education. This practice was inaugurated in the January, 1931, number when 587 titles were listed, accompanied by a topical index. Invitations to coöperate in the compilation of the list of Doctors' studies under way during the present school year have been sent to the following institutions: Boston, Chicago, California, Catholic University, Columbia, Cornell, George Washington, Harvard, Illinois, Indiana, Iowa, Johns Hopkins, Kansas, Kentucky, Michigan, Minnesota, Missouri, New York, Ohio, Oregon, Peabody, Pennsylvania, Pittsburgh, Southern California, Stanford, Texas, Washington, Wisconsin, and Yale.

If the reader is aware of other graduate institutions which have under way Doctors' theses in education, please send such information to the editor of this department, or feel free to have the institution forward a list of its Doctors' studies without a direct invitation from the editor.

1932 Yearbook on Character Education.—Character education is the subject of the 1932 yearbook of the Department of Superintendence. The commission which is preparing it is composed of the following people:

Bertie Backus, principal, Powell Junior High School, Washington, D. C.
George S. Counts, associate director, International Institute, Teachers College, Columbia University.
Frank N. Freeman, professor of educational psychology, University of Chicago.
C. B. Glenn, superintendent of schools, Birmingham, Alabama.
Paul T. Rankin, supervising director, department of research and adjustment, Board of Education, Detroit.
Belle M. Ryan, assistant superintendent of schools, Omaha, Nebraska.
Goodwin Watson, associate professor of education, Teachers College, Columbia University.
David E. Weglein, superintendent of schools, Baltimore, Maryland.
A. L. Threlkeld, superintendent of schools, Denver, Colorado, Chairman.

The commission is assisted by S. D. Shankland, executive secretary of the Department of Superintendence, John K. Norton and Margaret Alltucker Norton, until recently head and member respectively of the department of research, National Education Association.

The chapters of the yearbook, as planned to date, are as follows:

1. Analysis of American Life from the Viewpoint of Character Education
2. Agencies of Character Education
3. Objectives in Character Education
4. Theories of Character Education
5. Tests and Measurement in Character Education
6. Review of Research Studies in Character Education
7. How the Curriculum Content and Method May Contribute to Character Education

8. How Classroom Procedure May Contribute to Character Education
9. How the Life of the School May Contribute to Character Education
10. How Individual Counseling May Contribute to Character Education
11. The Teacher in Character Education
12. Organization of a School System in Relation to Character Education
13. The Administration of a School System in Relation to Character Education
14. Home and School Coöperation
15. The Relation of the School to Other Organizations and Institutions in Character Education

The purpose of the yearbook is to stimulate thought on the subject of character education rather than to attempt to present any specific program for the schools to follow. To make statements definite enough to accomplish this objective without seeming to place undue sanction upon specific theories, procedures, and the like, is a difficult task, yet this is what the commission is attempting.

Survey of the Dental Curriculum.—For some years many teachers of dentistry have felt the need for a systematic study and analysis of the dental curriculum. Curriculum content and its arrangement have been much discussed. Several national organizations have outlined curriculums or specified certain requirements for rating purposes, but no general agreement has been reached as to what constitutes adequate training for the practice of dentistry and how that training can best be given. Like the curriculums in most other professional fields, the dental curriculum has grown by accretion from time to time without thorough revision of the whole.

Among the forty-three dental schools of the United States and Canada, there is a considerable difference of opinion regarding the pre-dental requirements and the content and length of the dental curriculum. A number of different schemes for the undergraduate curriculum, such as the two-three plan, the one-four plan, and the two-four plan, are in operation, leading to embarrassment and confusion. The comprehensive Report on Dental Education made by the Carnegie Foundation for the Advancement of Teaching in 1926 criticized the dental curriculum and recommended the adoption of a requirement of at least two years of suitable pre-professional college work and a reorganization of the undergraduate dental curriculum, but it included no thorough analysis of curriculum content.

To correct these confusing conditions, the American Association of Dental Schools in 1929 decided to make a survey of the dental curriculum. The Association was granted $20,000 by the Carnegie Corporation for financing the survey. The following committee to have charge of the project was appointed: Wallace Seccombe, University of Toronto, Chairman; John T. O'Rourke, University of Louisville, Secretary-Treasurer; Arthur D. Black, Northwestern University; H. Edmund Friesell, University of Pittsburgh; and Harry M. Semans, Ohio State University. Floyd W. Reeves of the University of Chicago and W. W. Charters of Ohio State University were selected as technical advisers to the committee. L. E. Blauch has been appointed Executive Secretary to devote full time to the work of the survey. The office of the Executive Secretary is located in the Northwestern University Dental School.

The committee proposes to ascertain the requirements for dental health service and to determine the knowledge and skill which should be possessed by the dentist properly to render this service. The committee hopes it may be possible, at the conclusion of the study, to recom-

mend a curriculum which will prepare adequately the undergraduate for dental practice, including the development of character, an appreciation of the humanities, and a scientific outlook which will assure the graduate's interest in dental research. The study does not include curriculums for post-graduate work or for the training of specialists in dentistry. Rather, it is concerned with the training of general dental practitioners to render adequate dental health service in rural as well as urban centers. The findings of the committee will be reported to the American Association of Dental Schools previous to publication and will be subject to the approval of the Association.

Educational Research at Ohio State.—The Bureau of Educational Research at the Ohio State University, founded in 1922 by B. R. Buckingham and now under the direction of W. W. Charters, is composed of seven units.

Three of these seven units are devoted primarily to service, although various research projects also are carried on by them. These three service units are as follows:

1. The Reference Division, which contains the Reference Library, of which Mary Rebecca Lingenfelter is in charge.
2. The Editorial Division, under the direction of Josephine MacLatchy.
3. The Appointments Division, which is the University teacher placement agency, and is directed by Earl W. Anderson.

Research is the major function of the four remaining units of the Bureau, which are as follows:

1. The Division of Buildings and Grounds, directed by T. C. Holy, who also is associated with the Ohio State Department of Education.
2. The Division of Achievement Tests, under the direction of Ralph W. Tyler.
3. The Division of College Curriculum, in charge of Edgar Dale.
4. The Division of College Personnel, of which W. H. Cowley is director.

The senior staff listed above is aided by five research assistants and a junior staff consisting of clerical and secretarial workers. The Bureau occupies space in the Education Building, which also houses the administrative offices of the College of Education, offices of its various departments, and the Education Library.

Some of the more important projects may be listed. The Division of Buildings and Grounds assists school boards wishing to initiate building programs by making surveys for them and, in addition, conducts numerous school surveys for research purposes.

The Division of Achievement Tests is coöperating with some fourteen departments on the campus in an effort to improve existing types of tests and examinations. Enlargement of this program in the fields of botany, zoology, and chemistry has been made possible by a grant from the American Council of Education.

The College Curriculum Division is assisting in coöperative curriculum revision programs in the College of Veterinary Medicine and the College of Education. It also is coöperating with the State Director of Health and Physical Education in a study of pupil interests in the field of health.

The head of this Division and P. W. Holaday during this year are engaged in motion-picture research under the direction of the Committee on Educational Research of the Payne Fund of New York City. Dr. Holaday is making a study of the retention of motion-picture content by children, while Dr. Dale is analyzing the content, studying the attendance of children, and initiating a project designed

to raise children's standards of taste in motion pictures.

The College Personnel Division is working with other groups in an evaluation of freshman week, a survey of the quarter system as opposed to the semester system, and various studies regarding faculty contact with students.

Two annual conferences are directed by members of the Bureau staff—the Education Conference, under the direction of Dr. Cowley, and the Institute for Education by Radio, of which Dr. Lumley is the executive officer. In connection with the latter, a course in education by radio is given during the summer quarter. A recent addition to the Bureau is a radio unit, headed by Hillis Lumley, which has as its objective, research in this field.

The proceedings of the two conferences are published annually by the Editorial Division, directed by Miss MacLatchy. "Education on the Air," the proceedings of the 1930 Institute for Education by Radio, had the honor of being listed among the starred books in the sixty best educational publications of 1930. In addition to these two volumes, the Bureau publishes the *Journal of Higher Education* and the *Educational Research Bulletin*. A series of educational monographs also is issued, the latest contribution in this area being a study of the Bing Law.

Although most of the time of the staff members is devoted to research, each is teaching at least one course. Their total teaching load amounts to forty-five hours, the time of one and one-half full professors distributed among the eight staff members.

The Reference Library prepares bibliographies and places at the disposal of the Bureau staff a collection of research material in the field of education. It maintains a complete exhibit of current elementary- and secondary-school textbooks, and of achievement tests. This material is used not only by the Bureau, but by other research workers on the faculty, graduate students, and the schools of the state.

Organization of Educational Research at the University of Chicago.—Beginning with the Autumn Quarter, 1931, the Department of Education of the University of Chicago will give a new type of advanced research course. Instead of registering students for any research problem in which they are interested, each instructor has announced the problem on which he will be working constructively during the quarter in which the course is given. For example, W. C. Reavis will carry on intensive research concerning "The Committee Organization of Boards of Education;" G. T. Buswell will organize critical experiments on the constancy of eye movement habits and the relation of pheneloscope records of rhythm and pronunciation in reading to photographic records of eye movement; K. J. Holzinger will carry on advanced study in statistics with special application to the problem of the factor theories of mental abilities; and W. S. Gray will carry on a series of critical studies and investigations of the nature of the reading activities required in studying secondary-school subjects. The students who register for these and other similar courses will participate with the instructor in the various steps and stages of the studies which he is pursuing. It is hoped through participation in research, under the guidance of an expert in a field, students may acquire keener insight into the problems of research and make more rapid growth in acquiring efficiency in research activities.

Philadelphia Research.—Six years ago, upon recommendation of the superintend-

ent of schools, The Board of Public Education of Philadelphia created a Division of Educational Research. The functions of this division are to coördinate, extend, and direct research activities within the system to the end that reliable data may be available upon which to base the educational policies of the superintendent and his assistants. The division has gradually extended its activities to the point where professional stimulation and guidance have been afforded to all teachers and principals. The data gathered by the division have been utilized so as to bring about a more successful classification of pupils, and to introduce a diagnostic analysis of pupil performance upon which more effective teaching may be based. Analyses of age-grade status and of promotion rates within the system have led to administrative adjustments which result in more regular and satisfactory progress of pupils through the grades. Careful consideration of the wide variations in ability and achievements shown by the pupils of a given grade have directed attention to the importance of the adjustment of school requirements to the capacities and needs of pupils. Thus definite progress has been made toward individualized instruction and specific teaching.

The results of research activities throughout the system are recorded in the increased achievements of pupils revealed by the application of objective educational measurements. In a series of mimeographed and printed bulletins, the division has, from time to time, presented definite evidences of improvement in pupil achievement. These reports also direct attention to increased uniformity in interpretations of the several courses of study and to the desirability of remedial teaching in certain specific phases of school work. Thus, each year's progress has been built upon the achievements of the preceding year with both teachers and pupils growing more successful in their work as it comes to be based upon clearly defined goals and measured progress toward their achievement.

In addition to general research stimulation and guidance throughout the system, the division has pursued a number of projects which have furnished the superintendent with objective data upon which to make recommendations for the improvement of administrative policy. These studies have been concerned with such subjects as curriculum construction, teachers' salaries, teaching load in high schools, the efficiency of operation of certain school units, the measurable improvement of pupils in the command of oral and written English attained during the four-year high school period, and the desirable extent of individualization of instruction as represented in ability grouping, in individual assignment sheets, and in the establishment of special classes.

An important phase of the work of the division is that of objective test construction. It is desirable that such tests be definitely representative of the several courses of study in order that the results of their administration may direct the interpretation of these courses or indicate the need and character of revisions that may aid teachers and pupils in the more effective achievement of specific goals. An extensive range of tests is now available covering the requirements of most of the subjects in all Grades from I to XII.

A plan of standard scores has been developed which will aid teachers and principals in the interpretation of test scores and in the comparison of pupil performance on tests in different subjects and at different periods throughout the year. The division is now experimenting with a plan of "ability standards"

whereby an adjustment may be made for each pupil on a basis of his measured ability to perform in the work of the grade. In this way, it will be possible to bring about an individualization of instruction which will avoid the unfair competition of one pupil with another and substitute for it a more stimulating competition with one's own past achievement.

The division is continuing its long-term experimental study of the effects of ability grouping. Sufficient progress has been made in the control of conditions surrounding the study to warrant a two-year research program in an effort to secure objective measurement of the results of this plan of administration. Another experimental study is directed toward measuring the effects of the application of the new course of study in reading in Grades IV to VI. Still another project is the discovery and reporting of methods of teaching and drill employed in schools whose pupils show, for a given term, exceptionally good progress in the command of the fundamental operations in arithmetic. Extensive administrative studies are being made of teaching load in senior high schools and of the extent of provision for individual instruction.

The division staff is made up of a director, Philip A. Boyer, five professional assistants, and nine clerical assistants. Two of the professional assistants have the rank of special assistant. One of these directs attention solely to the field of secondary education, and the other to that of elementary education. Three assistants are ranked as supervisors. In addition to aiding in the research work of the division, these supervisors visit principals and give assistance in the construction and administration of group tests and in the analysis and interpretation of results. Upon the basis of these analyses, remedial procedures are developed that have specific application to the particular problems in hand. In this way, the division is endeavoring to recognize and render assistance in the solution of the peculiar problems of different local and school communities.

The division coöperates in the independent activities of the schools by making available as reference material the technical research data it has assembled. The staff of the division gives advice and assistance and on occasion coöperates in carrying out a project that seems to have important implications for improvement in some phase of school work. Increasing numbers of teachers and principals seek the resources and advice of the division as aids in solving teaching and supervisory problems. Thus, an important activity of the division concerns itself with supporting and stimulating the research attitude among teachers and principals by providing reference material and coöperating in individual research projects.

The projects of the division are varied in character and scope. Some activities, such as the experiment in ability grouping, extend over long terms. Others are continuous, such as the analysis and interpretation of promotion rates. Still others are periodic, such as the diagnosis of pupil difficulties at the beginning of a term or the survey of the results of instruction at the close of a term. In addition to these, there are incidental projects assigned by the superintendent, or suggested by the problems of individual teachers or principals. It is the practice of the division to make special report in printed or mimeographed form of the results of research projects that have wide application. These bulletins are made the subjects of professional discussions among superintendents, principals, and teachers.

California State Division of Research and Statistics.—The Division of Research and Statistics is concerned primarily with the administrative and service activities of the state department of education. It is essentially an adjunct of the office of the superintendent of public instruction, performing many functions legally allocated to this office, and having delegated to it many activities of administrative character with relation both to the state department and the development and execution of general policies for the state school system.

The functions of the Division, of which Walter E. Morgan is director, may be classified under four heads: administration, research and statistics, service, and coördination. The development of these functions in carrying on the activities of the Division in relation to the remainder of the state department of education and the public-school system determines the specific activities and the program for 1931.

Responsibility for such clerical-administrative activities as the apportionment of state school funds with prescription of record and report forms and blanks, auditing of annual reports, and preparation of statistical data for the departmental reports continue as the primary functions. In this capacity, the Division exercises a purely clerical-administrative capacity as an adjunct to the office of the superintendent of public instruction, with only an indirect relationship to the state school system.

The research and statistics function requires the continued development of techniques and methods for the presentation of significant data relative to the state school system, and calls for expanded interpretation of the data collected and presented, together with an extension of specialized and localized studies applied to selected types of problems.

The service functions will be expanded in order that more definite and specific services may be extended to the superintendent and to the state department, as well as to the school officers throughout the state school system. This phase of the Division's responsibility requires expansion more than any other at the present time.

The coördination function relates to both the activities of the several divisions of the state department and to relationships with school officers throughout the state. More numerous and more effective efforts can be made in providing closer coördination of the various activities which impinge upon those of the Division.

In 1931, the program involves specific projects under each of the four major functions.

Clerical-administrative projects are:

1. Continuance of regular activities, such as apportionment of school funds, auditing of reports, etc.
2. Continued simplification and coördination of all record and report forms.
3. Development of standardized record and report forms for the purpose of establishing greater uniformity in data collected. This includes such basic forms as the school district budget, school warrants, and classifications of school expenditures.
4. Standardization of fiscal procedures for schools, including the routing of school warrants and unified control of records of school funds.

These projected activities are primarily for the purpose of giving closer definition and uniformity to the basic records from which official reports are made to the office of the superintendent of public instruction. This will serve to improve the value of statistical data from the point of view of the state department as well as from that of the state school system.

Projected activities in the field of research and statistics include:

1. Continuance of present general and special tabulations of statistical data; coöperation with members of the department and school officers in the field in gathering and compiling specialized types of data; coöperation with candidates for higher degrees in gathering and compiling data in studies of state-wide significance; and interpretation of data as applied to localized or specialized situations.

2. Establishment of a card file of abstracts of research studies completed in California and elsewhere, the abstracts to be developed in coöperation with Phi Delta Kappa and to be made available to departmental staff members and to school officers upon request. These abstracts will be classified and filed in such a way as to make them readily serviceable.

3. Coöperation with city superintendents in establishing at Sacramento a research coördinator to serve city districts by bringing together results of administrative studies, coördinating research projects, and making available for the use of city superintendents composite reports of such studies. This will be developed in coöperation with the Association of Public School Superintendents.

4. Extension of endeavors to coördinate research activities throughout the state and to develop coöperative attack upon research problems.

5. Completion of numerous special studies relating to secondary schools for projected publications of the department.

6. Coöperation with staff members in preparing statistical summaries for reference purposes similar to the card summaries of data from the 1930 annual reports.

The program of service activities includes:

1. Coöperation with staff members in providing information, criticizing forms and questionnaires, etc.

2. Provision of a publication reference service whereby all publications coming to the division will be referred to staff members interested.

3. Coöperation with school officers by representing the state department in conferences, giving addresses, etc.

4. Conduct of special investigations on request, including surveys in the field.

5. Provide interpretations of data upon request from staff members or from officers in the field.

Increasing stress will be placed on the service function, which will be developed as the primary objective of the Division in terms of importance and value to both the department and the public-school system.

The coördination and integration function should embrace activities designed for the purpose of coördinating and integrating the activities within the department itself, relations between the department and the officers in the field, activities, programs, and policies throughout the state school system. Service on state-wide committees, direction given to committees and groups in the field, and interpretative service to staff members should comprise the major aspects of activities in this field. In conjunction with this program, there will be developed a program of publication which includes:

1. Publications at present established and being issued, such as the biennial report data, statistics of city school districts, junior-college statistics, directories, special analysis from October reports, and special studies in California schools.

2. Special studies, interpretative and statistical, applying to localized situations or specialized problems; e. g., comparative data for selected cities and high schools, or relating to such problems as teachers' salaries, school finance, etc.

3. Complete analysis of public-school finance.

4. Bulletin on public education in California.

5. A bulletin on California secondary education.

6. Gists of the state school code.

7. Application of an equalization-fund plan to the school districts of California.

An activity which might be classified under the function of coördination and integration, but which probably is related to all functions of the Division, is that of coöperation in the development of leg-

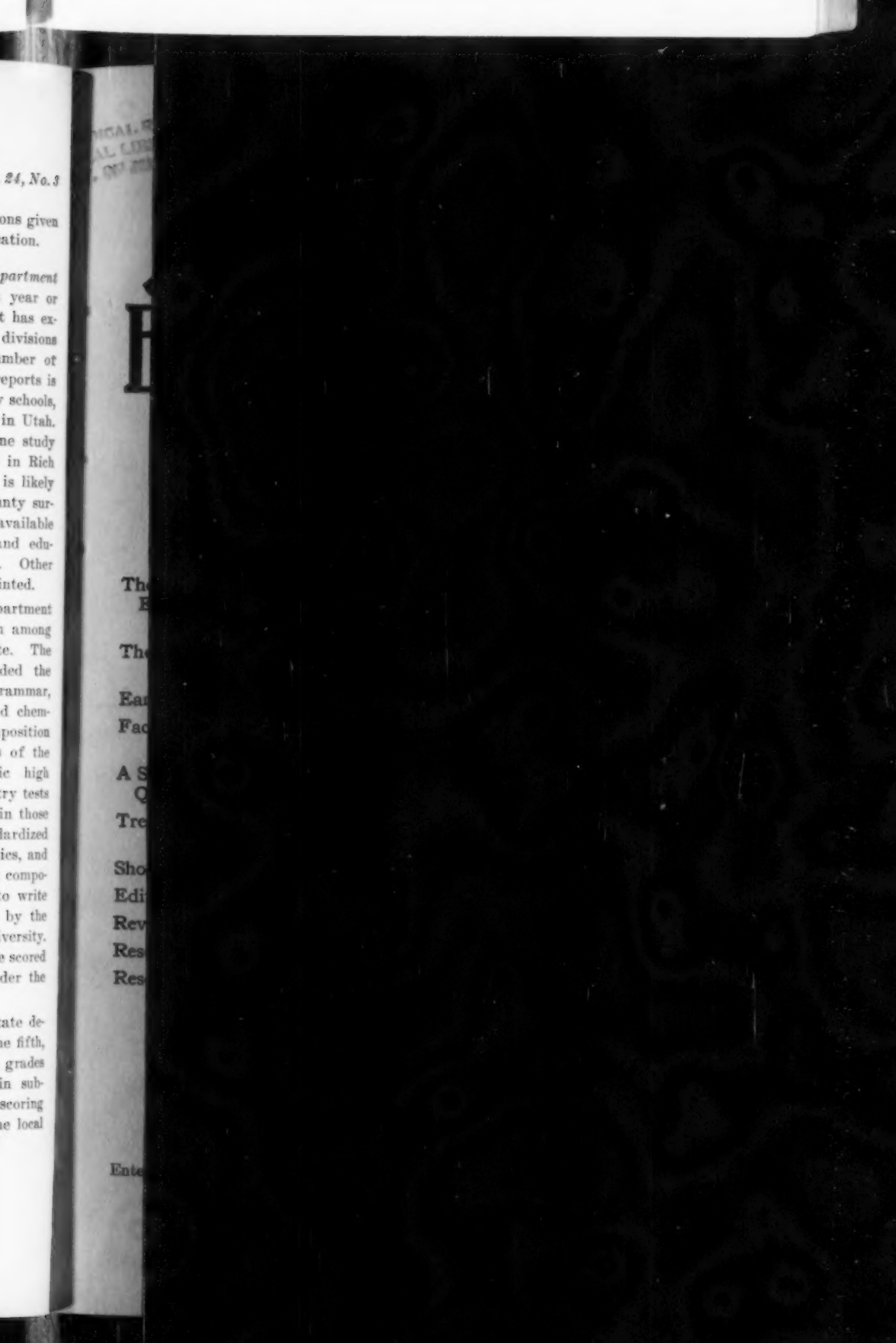
islative recommendations for the state department. Such recommendations are concerned with clarification of the school code and the drafting of bills relative to school finance and organization and administration. Related to this activity, and perhaps properly a part of the coördination function, is a project to develop throughout the state more uniform observance and interpretation by public-school officers of the state laws relating to public-school finance.

Bureau of Educational Measurements at the University of North Dakota.—The primary purpose of the Bureau is to give assistance in the measuring and interpreting of the results of teaching in the schools of the state, to collect and tabulate the results of educational tests in the schools, and to assist in any way possible in testing programs. During the past year, the Bureau has published a bulletin entitled "Educational Achievement in North Dakota City, Town, and Rural Schools." This bulletin contains the results of testing with the New Stanford Achievement Tests in the state. A few copies still are available for general distribution. During the next year it is planned to issue a similar bulletin which will give the results on other standardized tests which have been given in the schools.

Another project of the Bureau is the evaluation of workbooks in the various subject-matter fields, including first-year algebra and geometry.

The director of the Bureau, Robert D. Cole, has been working in close collaboration with the state department of education in the improvement of eighth-grade examinations for rural schools and in the preparation of tests in the various high-school subjects for unclassified schools. It is probable that he also will take over the supervision of the examinations given by the state department of education.

Research in the Utah State Department of Education.—During the last year or two, the Utah State Department has extended the service of its several divisions to include the survey of a number of school districts. One of these reports is a survey of the Duchesne County schools, which is a typical rural district in Utah. A survey similar to the Duchesne study is being conducted by the staff in Rich county at the present time. It is likely that the report of the Rich County survey will be mimeographed and available for interested school systems and educators during the current year. Other smaller surveys have not been printed.

During recent years, the department has conducted a testing program among all the high schools of the state. The testing program in 1930 included the following subjects: English grammar, English composition, physics, and chemistry. The grammar and composition tests were given to all members of the graduating class in the public high schools. The physics and chemistry tests were given to all pupils enrolled in those subjects in May, 1930. Standardized tests were used in grammar, physics, and chemistry. In measuring English composition, all seniors were required to write compositions which were graded by the English faculty of the state university. The tests in the other subjects were scored in the state department office under the direction of the staff.

During the recent years, the state department also has given tests to the fifth, sixth, seventh, eighth, and ninth grades throughout the state in the main subjects taught in these grades. The scoring of the papers has been done by the local school people.